

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
5 October 2000 (05.10.2000)

PCT

(10) International Publication Number
WO 00/58473 A3

(51) International Patent Classification⁷: C12N 15/12,
C07K 14/47, 16/18, G01N 33/566, C12Q 1/68, C12N
15/11, 15/62, A01K 67/027, A61K 38/00

Richard, A. [US/US]; 191 Leete Street, West Haven,
CT 06516 (US). LEACH, Martin [GB/US]; 884 School
Street, Webster, MA 01570 (US).

(21) International Application Number: PCT/US00/08621

(74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Cohn, Ferris,
Glovsky and Popeo, P.C., One Financial Center, Boston,
MA 02111 (US).

(22) International Filing Date: 31 March 2000 (31.03.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/127,607 31 March 1999 (31.03.1999) US
60/127,636 2 April 1999 (02.04.1999) US
60/127,728 5 April 1999 (05.04.1999) US
09/540,763 30 March 2000 (30.03.2000) US

(81) Designated States (*national*): AE, AL, AM, AT, AU, AZ,
BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK,
DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL,
IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU,
LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT,
RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA,
UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent
(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent
(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU,
MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM,
GA, GN, GW, ML, MR, NE, SN, TD, TG).

(63) Related by continuation (CON) or continuation-in-part
(CIP) to earlier applications:

US 60/127,607 (CIP)
Filed on 31 March 1999 (31.03.1999)
US 60/127,636 (CIP)
Filed on 2 April 1999 (02.04.1999)
US 60/127,728 (CIP)
Filed on 5 April 1999 (05.04.1999)
US 09/540,763 (CIP)
Filed on 30 March 2000 (30.03.2000)

Published:

— With international search report.
— Before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments.

(71) Applicant (*for all designated States except US*): CURA-
GEN CORPORATION [US/US]; 555 Long Wharf Drive,
11th Floor, New Haven, CT 06511 (US).

(88) Date of publication of the international search report:
25 January 2001

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): SHIMKETS,

*For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.*

WO 00/58473 A3

(54) Title: NUCLEIC ACIDS INCLUDING OPEN READING FRAMES ENCODING POLYPEPTIDES; "ORFX"

(57) Abstract: The present invention provides open reading frames encoding isolated polypeptides, as well as polynucleotides encoding ORFX and antibodies that immunospecifically bind to ORFX or any derivative, variant, mutant, or fragment of the ORFX polypeptides, polynucleotides or antibodies. The invention additionally provides methods in which the ORFX polypeptide, polynucleotide and antibody are used in detection and treatment of a broad range of pathological states, as well as to other uses.

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/US 00/08621

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/12 C07K14/47 C07K16/18 G01N33/566 C12Q1/68
 C12N15/11 C12N15/62 A01K67/027 A61K38/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K G01N A01K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

BIOSIS, EMBASE, MEDLINE, CAB Data, PAJ, EP0-Internal, WPI Data, STRAND

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	COLE S.T.: "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence." NATURE, vol. 393, 11 June 1998 (1998-06-11), XP002144873 sequence	
A	--- LAMERDIN J.E.: "Sequence analysis of a 3.5 Mb contig in human 19p13.3 containing a serine protease gene cluster." EMEST DATABASE ENTRY, 8 February 1999 (1999-02-08), XP002144874 sequence --- -/--	

☒ Further documents are listed in the continuation of box C.☐ Patent family members are listed in annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

21 August 2000

Date of mailing of the international search report

23.11.00

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
 Fax: (+31-70) 340-3016

Authorized officer

Hix, R.

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 00/08621

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	<p>M.D. ADAMS ET AL.: "The genome sequence of <i>Drosophila melanogaster</i>." SCIENCE, vol. 287, 24 March 2000 (2000-03-24), pages 2185-2195, XP002144875 the whole document -----</p>	6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 00/08621

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 27 to 32 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

claims 1 to 32 partially

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claim : 1 to 32 partially

Isolated nucleic acid molecule encoding a polypeptide comprising an amino acid sequence that is at least 85% identical to a polypeptide including an amino acid sequence selected from a group consisting of SEQ ID NO 2n wherein n is 1, oligonucleotides less than 100 nucleotides in length and comprising at least 6 contiguous nucleotides from the above sequence, polypeptides encoded by said nucleotides, antibodies that bind to said polypeptide, pharmaceutical composition comprising said polypeptide and methods of detection, screening, therapeutic uses involving said polypeptide.

2. Claim : .

Inventions 2 to 3161

claims 1 to 32 partially :

Isolated nucleic acid molecule encoding a polypeptide comprising an amino acid sequence that is at least 85% identical to a polypeptide including an amino acid sequence selected from a group consisting of SEQ ID NO 2n wherein n is 2 to 3161, oligonucleotides less than 100 nucleotides in length and comprising at least 6 contiguous nucleotides from the above sequence, polypeptides encoded by said nucleotides, antibodies that bind to said polypeptide, pharmaceutical composition comprising said polypeptide and methods of detection, screening, therapeutic uses involving said polypeptide.



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C12N 15/12, C07K 14/47, 16/18, G01N 33/566, C12Q 1/68, C12N 15/11, 15/62, A01K 67/027, A61K 38/00	A2	(11) International Publication Number: WO 00/58473 (43) International Publication Date: 5 October 2000 (05.10.00)																												
(21) International Application Number: PCT/US00/08621 (22) International Filing Date: 31 March 2000 (31.03.00) (30) Priority Data: <table border="0"> <tr> <td>60/127,607</td> <td>31 March 1999 (31.03.99)</td> <td>US</td> </tr> <tr> <td>60/127,636</td> <td>2 April 1999 (02.04.99)</td> <td>US</td> </tr> <tr> <td>60/127,728</td> <td>5 April 1999 (05.04.99)</td> <td>US</td> </tr> <tr> <td>09/540,763</td> <td>30 March 2000 (30.03.00)</td> <td>US</td> </tr> </table> (63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications <table border="0"> <tr> <td>US</td> <td>60/127,607 (CIP)</td> </tr> <tr> <td>Filed on</td> <td>31 March 1999 (31.03.99)</td> </tr> <tr> <td>US</td> <td>60/127,636 (CIP)</td> </tr> <tr> <td>Filed on</td> <td>2 April 1999 (02.04.99)</td> </tr> <tr> <td>US</td> <td>60/127,728 (CIP)</td> </tr> <tr> <td>Filed on</td> <td>5 April 1999 (05.04.99)</td> </tr> <tr> <td>US</td> <td>09/540,763 (CIP)</td> </tr> <tr> <td>Filed on</td> <td>30 March 2000 (30.03.00)</td> </tr> </table> (71) Applicant (for all designated States except US): CURAGEN CORPORATION [US/US]; 555 Long Wharf Drive, 11th Floor, New Haven, CT 06511 (US).		60/127,607	31 March 1999 (31.03.99)	US	60/127,636	2 April 1999 (02.04.99)	US	60/127,728	5 April 1999 (05.04.99)	US	09/540,763	30 March 2000 (30.03.00)	US	US	60/127,607 (CIP)	Filed on	31 March 1999 (31.03.99)	US	60/127,636 (CIP)	Filed on	2 April 1999 (02.04.99)	US	60/127,728 (CIP)	Filed on	5 April 1999 (05.04.99)	US	09/540,763 (CIP)	Filed on	30 March 2000 (30.03.00)	(72) Inventors; and (75) Inventors/Applicants (for US only): SHIMKETS, Richard, A. [US/US]; 191 Leete Street, West Haven, CT 06516 (US). LEACH, Martin [GB/US]; 884 School Street, Webster, MA 01570 (US). (74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Cohn, Ferris, Glovsky and Popeo, P.C., One Financial Center, Boston, MA 02111 (US). (81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>Without international search report and to be republished upon receipt of that report.</i>
60/127,607	31 March 1999 (31.03.99)	US																												
60/127,636	2 April 1999 (02.04.99)	US																												
60/127,728	5 April 1999 (05.04.99)	US																												
09/540,763	30 March 2000 (30.03.00)	US																												
US	60/127,607 (CIP)																													
Filed on	31 March 1999 (31.03.99)																													
US	60/127,636 (CIP)																													
Filed on	2 April 1999 (02.04.99)																													
US	60/127,728 (CIP)																													
Filed on	5 April 1999 (05.04.99)																													
US	09/540,763 (CIP)																													
Filed on	30 March 2000 (30.03.00)																													
(54) Title: NUCLEIC ACIDS INCLUDING OPEN READING FRAMES ENCODING POLYPEPTIDES; "ORFX" (57) Abstract <p>The present invention provides open reading frames ORFX, encoding isolated polypeptides, as well as polynucleotides encoding ORFX and antibodies that immunospecifically bind to ORFX or any derivative, variant, mutant, or fragment of the ORFX polypeptides, polynucleotides or antibodies. The invention additionally provides methods in which the ORFX polypeptide, polynucleotide and antibody are used in detection and treatment of a broad range of pathological states, as well as to other uses.</p>																														

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY

5

BACKGROUND OF THE INVENTION

The invention relates generally to nucleic acids and polypeptides encoded thereby, and methods of using these nucleic acids and polypeptides.

10

SUMMARY OF THE INVENTION

The invention is based in part on the discovery of nucleic acids that include open reading frames encoding novel polypeptides, and on the polypeptides encoded thereby. The nucleic acids and polypeptides are collectively referred to herein as "ORFX".

15

Accordingly, in one aspect, the invention provides an isolated nucleic acid molecule (SEQ ID NO:2*n*-1, wherein *n* is an integer between 1-3161), that encodes novel polypeptide, or a fragment, homolog, analog or derivative thereof. The nucleic acid can include, *e.g.*, a nucleic acid sequence encoding a polypeptide at least 85% identical to a polypeptide comprising the amino acid sequences of SEQ ID NO:2*n*, wherein *n* is an integer between 1-3161. The nucleic acid can be, *e.g.*, a genomic DNA fragment, or a cDNA molecule.

20

Also included in the invention is a vector containing one or more of the nucleic acids described herein, and a cell containing the vectors or nucleic acids described herein.

The invention is also directed to host cells transformed with a recombinant expression vector comprising any of the nucleic acid molecules described above.

25

In another aspect, the invention includes a pharmaceutical composition that includes an ORFX nucleic acid and a pharmaceutically acceptable carrier or diluent.

In a further aspect, the invention includes a substantially purified ORF polypeptide, *e.g.*, any of the ORFX polypeptides encoded by an ORFX nucleic acid, and fragments, homologs, analogs, and derivatives thereof. The invention also includes a pharmaceutical composition that includes a ORFX polypeptide and a pharmaceutically acceptable carrier or diluent.

5 In a still a further aspect, the invention provides an antibody that binds specifically to an ORFX polypeptide. The antibody can be, *e.g.*, a monoclonal or polyclonal antibody, and fragments, homologs, analogs, and derivatives thereof. The invention also includes a pharmaceutical composition including ORFX antibody and a pharmaceutically acceptable carrier or diluent. The invention is also directed to isolated antibodies that bind to an epitope on a
10 polypeptide encoded by any of the nucleic acid molecules described above.

The invention also includes kits comprising any of the pharmaceutical compositions described above.

The invention further provides a method for producing an ORFX polypeptide by providing a cell containing a ORFX nucleic acid, *e.g.*, a vector that includes a ORFX nucleic
15 acid, and culturing the cell under conditions sufficient to express the ORFX polypeptide encoded by the nucleic acid. The expressed ORFX polypeptide is then recovered from the cell. Preferably, the cell produces little or no endogenous ORFX polypeptide. The cell can be, *e.g.*, a prokaryotic cell or eukaryotic cell.

The invention is also directed to methods of identifying an ORFX polypeptide or nucleic
20 acids in a sample by contacting the sample with a compound that specifically binds to the polypeptide or nucleic acid, and detecting complex formation, if present.

The invention further provides methods of identifying a compound that modulates the activity of a ORFX polypeptide by contacting ORFX polypeptide with a compound and determining whether the ORFX polypeptide activity is modified.

25 The invention is also directed to compounds that modulate ORFX polypeptide activity identified by contacting a ORFX polypeptide with the compound and determining whether the compound modifies activity of the ORFX polypeptide, binds to the ORFX polypeptide, or binds to a nucleic acid molecule encoding a ORFX polypeptide.

In a another aspect, the invention provides a method of determining the presence of or
30 predisposition of an ORFX-associated disorder in a subject. The method includes providing a sample from the subject and measuring the amount of ORFX polypeptide in the subject sample.

The amount of ORFX polypeptide in the subject sample is then compared to the amount of ORFX polypeptide in a control sample. An alteration in the amount of ORFX polypeptide in the subject protein sample relative to the amount of ORFX polypeptide in the control protein sample indicates the subject has a tissue proliferation-associated condition. A control sample is
5 preferably taken from a matched individual, *i.e.*, an individual of similar age, sex, or other general condition but who is not suspected of having a tissue proliferation-associated condition. Alternatively, the control sample may be taken from the subject at a time when the subject is not suspected of having a tissue proliferation-associated disorder. In some embodiments, the ORFX is detected using a ORFX antibody.

10 In a further aspect, the invention provides a method of determining the presence of or predisposition of an ORFX-associated disorder in a subject. The method includes providing a nucleic acid sample, *e.g.*, RNA or DNA, or both, from the subject and measuring the amount of the ORFX nucleic acid in the subject nucleic acid sample. The amount of ORFX nucleic acid sample in the subject nucleic acid is then compared to the amount of an ORFX nucleic acid in a
15 control sample. An alteration in the amount of ORFX nucleic acid in the sample relative to the amount of ORFX in the control sample indicates the subject has a tissue proliferation-associated disorder.

In a still further aspect, the invention provides method of treating or preventing or delaying a ORFX-associated disorder. The method includes administering to a subject in which
20 such treatment or prevention or delay is desired a ORFX nucleic acid, a ORFX polypeptide, or an ORFX antibody in an amount sufficient to treat, prevent, or delay a tissue proliferation-associated disorder in the subject.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention
25 belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In the case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and
30 examples are illustrative only and not intended to be limiting.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

DETAILED DESCRIPTION OF THE INVENTION

The invention provides novel polypeptides and nucleotides encoded thereby. The polynucleotides and their encoded polypeptides can be grouped according to the functions played by their gene products. Such functions include, structural proteins, proteins from which associated with metabolic pathways fatty acid metabolism, glycolysis, intermediary metabolism, calcium metabolism, proteases, and amino acid metabolism, etc.

Included in the invention are 3161 novel nucleic acid sequences and their encoded polypeptides. The sequences are collectively referred to as "ORFX nucleic acids" or ORFX polynucleotides" and the corresponding encoded polypeptide is referred to as a "ORFX polypeptide" or ORFX protein". For example, an ORFX nucleic acid according to the invention is a nucleic acid including an ORF1 nucleic acid, and an ORF polypeptide according to the invention is a polypeptide that includes the amino acid sequence of an ORF1 polypeptide. Unless indicated otherwise, "ORFX" is meant to refer to any of the ORF1-3161 sequences disclosed herein.

Table 1 provides a summary of the ORFX nucleic acids and their encoded polypeptides are summarized in Table 1. Nucleic acid sequences and polypeptide sequences for ORFX nucleic acids according to the invention is provided in the section of the specification entitled "Disclosed Sequences of ORFX Nucleic Acid and Polypeptide Sequences."

Column 1 of Table 1, entitled "ORF #", denotes an ORF number assigned to a nucleic acid containing an open reading frame according to the invention.

Column 2 of Table 1, entitled "Internal Identification number (Nucleic Acid Sequence Identification Number, Polypeptide Sequence Identification Number), provides an internal identification number for the indicated ORF, along with sequence identification numbers (SEQ ID NOs.) corresponding to the indicated ORF. In general, for an ORF_n according to the invention (wherein *n* is any integer from 1 to 3161), a nucleic acid corresponding to the ORF is SEQ ID NO:2*n*-1, and an amino acid sequence encoded by the ORF is SEQ ID NO:2*n*. For example, a nucleic acid sequence corresponding to an ORF1 nucleic acid is SEQ ID NO:1, and a polypeptide sequence corresponding to an ORF1 polypeptide is SEQ ID NO:2. Similarly, a

nucleic acid sequence corresponding to an ORF4 nucleic acid is SEQ ID NO:7, and a polypeptide sequence corresponding to an ORF4 polypeptide is SEQ ID NO:8; a nucleic acid sequence corresponding to an ORF198 nucleic acid sequence is SEQ ID NO:395, and a polypeptide sequence corresponding to an ORF198 polypeptide is SEQ ID NO:396. Nucleic acid sequences and polypeptide sequences for ORFX nucleic acids according to the invention are provided in the section of the specification entitled "Disclosed Sequences of ORFX Nucleic Acid and Polypeptide Sequences."

Column 2 of Table 1, entitled "Protein Similarity", lists previously described proteins that are related to polypeptides encoded by the ORFs. Genbank identifiers for the previously described proteins are provided. These can be retrieved from <http://www.ncbi.nlm.nih.gov/>.

To determine similarity to previously described proteins, polypeptides encoded by ORFX DNA sequences were tested using the Framesearch Algorithm against a nonredundant version of the GenPept Database from NCBI/Genbank. DNA sequences that had a score of '90' or above (Framesearch algorithm score, Edelman et. al. GCG Genetics) to a known protein were selected.

Open reading frames were extended beyond the region of the protein matched using standard DNA translation and codon tables. Novel proteins that lacked a protein match were translated against the standard genetic codons and proteins with an ORF at least 80 amino acids and containing a Methionine start are included in the Table.

Column 3 of Table 3, entitled "Protein Domains", lists previously described protein domains, designated by pfam entries, that are present in polypeptides encoded by the ORFs. Also included in column 3 are proteins in which these domains are present. The pfam entries can be retrieved from <http://pfam.wustl.edu/>. DNA sequences were translated in all six frames and tested using the Hmmer Algorithm against the Pfam Database (References to the algorithm and Pfam database can be found at <http://pfam.wustl.edu/>). Translated DNA sequences that matched a protein domain entry in the Pfam database AND had a score of 7.5' were selected.

Column 4 of Table 3, entitled "Protein Classification", lists the type of classification assigned for the protein, based on its homology. Examples of proteins in the classification include the following proteins:

Amylases

Amylase is responsible for endohydrolysis of 1,4-alpha-glucosidic linkages in oligosaccharides and polysaccharides. Variations in amylase gene may be indicative of delayed maturation and of various amylase producing neoplasms and carcinomas.

5 Amyloid

The serum amyloid A (SAA) proteins comprise a family of vertebrate proteins that associate predominantly with high density lipoproteins (HDL). The synthesis of certain members of the family is greatly increased in inflammation. Prolonged elevation of plasma SAA levels, as in chronic inflammation, results in a pathological condition, called amyloidosis, which affects the liver, kidney and spleen and which is characterized by the highly insoluble accumulation of SAA in these tissues. Amyloid selectively inhibits insulin-stimulated glucose utilization and glycogen deposition in muscle, while not affecting adipocyte glucose metabolism. Deposition of fibrillar amyloid proteins intraneuronally, as neurofibrillary tangles, extracellularly, as plaques and in blood vessels, is characteristic of both Alzheimer's disease and aged Down's syndrome.

15 Amyloid deposition is also associated with type II diabetes mellitus.

Angiopoeitin

Members of the angiopoietin/fibrinogen family have been shown to stimulate the generation of new blood vessels, inhibit the generation of new blood vessels, and perform several roles in blood clotting. This generation of new blood vessels, called angiogenesis, is also an essential step in tumor growth in order for the tumor to get the blood supply it needs to expand.

20 Variation in these genes may be predictive of any form of heart disease, numerous blood clotting disorders, stroke, hypertension and predisposition to tumor formation and metastasis. In particular, these variants may be predictive of the response to various antihypertensive drugs and chemotherapeutic and anti-tumor agents.

25 Apoptosis-related proteins

Active cell suicide (apoptosis) is induced by events such as growth factor withdrawal and toxins. It is controlled by regulators, which have either an inhibitory effect on programmed cell

death (anti-apoptotic) or block the protective effect of inhibitors (pro-apoptotic). Many viruses have found a way of countering defensive apoptosis by encoding their own anti-apoptosis genes preventing their target-cells from dying too soon. Variants of apoptosis related genes may be useful in formulation of anti-aging drugs.

Cadherin, Cyclin, Polymerase, Oncogenes, Histones, Kinases

Members of the cell division/cell cycle pathways such as cyclins, many transcription factors and kinases, DNA polymerases, histones, helicases and other oncogenes play a critical role in carcinogenesis where the uncontrolled proliferation of cells leads to tumor formation and eventually metastasis. Variation in these genes may be predictive of predisposition to any form of cancer, from increased risk of tumor formation to increased rate of metastasis. In particular, these variants may be predictive of the response to various chemotherapeutic and anti-tumor agents.

Colony-stimulating factor-related proteins

Granulocyte/macrophage colony-stimulating factors are cytokines that act in hematopoiesis by controlling the production, differentiation, and function of 2 related white cell populations of the blood, the granulocytes and the monocytes-macrophages.

Complement-related proteins

Complement proteins are immune associated cytotoxic agents, acting in a chain reaction to exterminate target cells to that were opsonized (primed) with antibodies, by forming a membrane attack complex (MAC). The mechanism of killing is by opening pores in the target cell membrane. Variations in 20 complement genes or their inhibitors are associated with many autoimmune disorders. Modified serum levels of complement products cause edemas of various tissues, lupus (SLE), vasculitis, glomerulonephritis, renal failure, hemolytic anemia, thrombocytopenia, and arthritis. They interfere with mechanisms of ADCC (antibody dependent cell cytotoxicity), severely impair immune competence and reduce phagocytic ability. Variants of complement genes may also be indicative of type I diabetes mellitus, meningitis neurological disorders such as nemaline myopathy, neonatal hypotonia, muscular disorders such as congenital myopathy and other diseases.

Cytochrome

The respiratory chain is a key biochemical pathway which is essential to all aerobic cells. There are five different cytochromes involved in the chain. These are heme bound proteins which serve as electron carriers. Modifications in these genes may be predictive of ataxia areflexia, dementia and myopathic and neuropathic changes in muscles. Also, association with various types of solid tumors.

Kinesins

Kinesins are tubulin molecular motors that function to transport organelles within cells and to move chromosomes along microtubules during cell division. Modifications of these genes may be indicative of neurological disorders such as Pick disease of the brain, tuberous sclerosis.

Cytokines, Interferon, Interleukin

Members of the cytokine families are known for their potent ability to stimulate cell growth and division even at low concentrations. Cytokines such as erythropoietin are cell-specific in their growth stimulation; erythropoietin is useful for the stimulation of the proliferation of erythroblasts. Variants in cytokines may be predictive for a wide variety of diseases, including cancer predisposition.

G-protein coupled receptors

G-protein coupled receptors (also called R7G) are an extensive group of hormones, neurotransmitters, odorants and light receptors which transduce extracellular signals by interaction with guanine nucleotide-binding (G) proteins. Alterations in genes coding for G-coupled proteins may be involved in and indicative of a vast number of physiological conditions. These include blood pressure regulation, renal dysfunctions, male infertility, dopamine associated cognitive, emotional, and endocrine functions, hypercalcemia, chondrodysplasia and osteoporosis, pseudohypoparathyroidism, growth retardation and dwarfism.

Thioesterases

Eukaryotic thiol proteases are a family of proteolytic enzymes which contain an active site cysteine. Catalysis proceeds through a thioester intermediate and is facilitated by a nearby histidine side chain; an asparagine completes the essential catalytic triad. Variants of thioester associated genes may be predictive of neuronal disorders and mental illnesses such as Ceroid Lipoffiscinosis, Neuronal 1, Infantile, Santavuori disease and more.

The key to the molecule type is as follows:

10	Abbrev:	Title:
	amylase	amylase protein
	amylaseinhib	amylase inhibitor
	amyloid	amyloid protein
15	apoptosis	apoptosis associated protein
	apoptosisinhib	apoptosis inhibitors
	apoptosisrecep	apoptosis receptors
	ATPase_associated	ATPase associated protein
	biotindep	biotin dependent enzyme/protein
20	cadherin	cadherin protein
	calcium_channel	calcium channel protein
	carboxylase	carboxylase protein
	cathepsin	cathepsin/carboxypeptidases
	cathepsininhib	cathepsin/carboxypeptidase inhibitor
25	chloride_channel	chloride channel protein
	collagen	collagen
	complement	complement protein
	complementrecept	complement receptor protein
	complementinhib	complement inhibitor
30	csf	colony stimulating factor
	csfrecept	colony stimulating factor receptor
	cyclin	cyclin protein
	cyto450	cytochrome p450 protein
	cytochrome	cytochrome related protein
35	deaminase	deaminase
	dehydrogenase	dehydrogenase
	desaturase	desaturase
	dna_rna_bind	DNA/RNA binding protein/factor
	dna_rna_inhib	DNA/RNA binding protein/factor inhibitor
40	dynein	dynein

	elastase	elastase
	elastaseinhib	elastase inhibitor
	eph	EPH family of tyrosine kinases
	esterase	esterase
5	esteraseinhib	esterase inhibitor
	fgf	fibroblast growth factor
	fgfreceptor	fibroblast growth factor receptor
	gaba	GABA receptor
	glucoamylase	glucoamylase
10	glucoronidase	glucoronidase
	glycoprotein	glycoprotein
	Guanylyl	guanylylate cyclase
	helicase	helicase
	histone	histone
15	HOM	homologous
	homeobox	homeobox protein
	hydrolase	hydrolase
	hydroxysteroid	hydroxysteroid associated protein
	hypoxanthine	hypoxanthine associated protein
20	immunoglob	immunoglobulin
	immunoglobrecept	immunoglobulin receptor
	interferon	interferon
	interleukin	interleukin
	interleukinrecept	interleukin receptor
25	isomerase	isomerase
	isomeraseinhibitor	isomerase inhibitor
	isomerasereceptor	isomerase receptor
	kinase	kinase
	kinaseinhibitor	kinase inhibitor
30	kinasereceptor	kinase receptor
	kinesin	kinesin
	laminin	laminin associated protein
	lipase	lipase
	metallothionein	metallothionein
35	MHC	major histocompatibility complex
	misc_channel	miscellaneous channel
	ngf	nerve growth factor
	nuci_recpt	nuclear receptor
	nuclease	nuclease
40	oncogene	oncogene associated protein
	oxidase	oxidase
	oxygenase	oxygenase
	peptidase	peptidase
	peroxidase	peroxidase
45	phosphatase	phosphatase
	phosphataseinhib	phosphatase inhibitor

	phosphorylase PIR	phosphorylase PIR DATABASE (release 56, 29-OCT-1998)
5	polymerase potassium_channel prostaglandin protease proteaseinhib reductase	polymerase potassium channel protein prostaglandin protease protease inhibitor reductase
10	ribosomalprot RTR	ribosomal associated protein EMBLDATABASE translated entries not to be incorporated into SWISS-PROT (20-JUL-1998)
15	SIM SPTR	similar EMBL DATABASE translated entries to be incorporated into SWISS-PROT (20-JUL-1998)
20	struct sulfotransferase SWP	structural associated protein sulfotransferase SWISS-PROT DATABASE (release 18-OCT-1998)
25	SWPN synthase tgf tgfreceptor thioesterase thiolase tm7	SWISS-PROT Update (release 11-NOV-98) synthase transforming growth factor transforming growth factor receptor thioesterase thiolase
30	tnf traffic tnfreceptor TRN	seven transmembrane domain G-protein coupled receptor necrosis factor receptor tumor necrosis factor tumor trafficking associated protein EMBL DATABASE translated entries update (20-JUL-1998)
35	transcriptfactor transferase transport tubulin ubiquitin	transcription factor transferase transport protein tubulin ubiquitin
40	unclassified	Protein not categorized into one of the aforementioned protein families
	water channel	water channel protein

Column 5 of Table 1, entitled, "Cells or Tissues in Which Gene is Expressed", denotes tissues, represented by five digit numbers, in which RNA homologous to the ORF nucleic acid sequences is present. Tissues or cells corresponding to the numbers are provided in Table 2.

ORFX nucleic acids, and their encoded polypeptides, according to the invention are useful in a variety of applications and contexts. For example, various ORFX nucleic acids and polypeptides according to the invention are useful, *inter alia*, as novel members of the protein families indicated in Table 1, and/or according to the presence of domains and sequence relatedness to previously described proteins as summarized in Table 1.

ORFX nucleic acids and polypeptides according to the invention can also be used to identify cell types listed in Table 1 for an indicated ORFX according to the invention. Additional utilities for ORFX nucleic acids and polypeptides according to the invention are disclosed herein.

ORFX Nucleic Acids

The novel nucleic acids of the invention include those that encode an ORFX or ORFX-like protein, or biologically active portions thereof. The nucleic acids include nucleic acids encoding polypeptides that include the amino acid sequence of one or more of SEQ ID NO:2n, wherein $n = 1$ to 3161. The encoded polypeptides can thus include, *e.g.*, the amino acid sequences of SEQ ID NO: 2, 4, 6, 8, 10, ..., 6310, 6312, 6314, 6316, 6318, 6320, and/or 6322.

In some embodiments, a nucleic acid encoding a polypeptide having the amino acid sequence of one or more of SEQ ID NO:2n (wherein $n = 1$ to 3161) includes the nucleic acid sequence of any of SEQ ID NO:2n-1 (wherein $n = 1$ to 3161), or a fragment thereof. Additionally, the invention includes mutant or variant nucleic acids of any of SEQ ID NO:2n-1 (wherein $n = 1$ to 3161), or a fragment thereof, any of whose bases may be changed from the disclosed sequence while still encoding a protein that maintains its ORFX-like activities and physiological functions. The invention further includes the complement of the nucleic acid sequence of any of SEQ ID NO:2n-1 (wherein $n = 1$ to 3161), including fragments, derivatives,

analogues and homolog thereof. The invention additionally includes nucleic acids or nucleic acid fragments, or complements thereto, whose structures include chemical modifications.

Also included are nucleic acid fragments sufficient for use as hybridization probes to identify ORFX-encoding nucleic acids (*e.g.*, ORFX mRNA) and fragments for use as polymerase chain reaction (PCR) primers for the amplification or mutation of ORFX nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA), RNA molecules (*e.g.*, mRNA), analogs of the DNA or RNA generated using nucleotide analogs, and derivatives, fragments and homologs thereof. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

"Probes" refer to nucleic acid sequences of variable length, preferably between at least about 10 nucleotides (nt), 100 nt, or as many as about, *e.g.*, 6,000 nt, depending on use. Probes are used in the detection of identical, similar, or complementary nucleic acid sequences. Longer length probes are usually obtained from a natural or recombinant source, are highly specific and much slower to hybridize than oligomers. Probes may be single- or double-stranded and designed to have specificity in PCR, membrane-based hybridization technologies, or ELISA-like technologies.

An "isolated" nucleic acid molecule is one that is separated from other nucleic acid molecules that are present in the natural source of the nucleic acid. Examples of isolated nucleic acid molecules include, but are not limited to, recombinant DNA molecules contained in a vector, recombinant DNA molecules maintained in a heterologous host cell, partially or substantially purified nucleic acid molecules, and synthetic DNA or RNA molecules. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated ORFX nucleic acid molecule can contain less than about 50 kb, 25 kb, 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material or culture medium when produced by recombinant techniques, or of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or a complement of any of this nucleotide sequence, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the nucleic acid sequence of any of SEQ ID NO:2*n*-1 (wherein *n*=1 to 3161) as a hybridization probe, ORFX nucleic acid sequences can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook *et al.*, eds., MOLECULAR CLONING: A LABORATORY MANUAL 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; and Ausubel, *et al.*, eds., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993.)

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to ORFX nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

As used herein, the term "oligonucleotide" refers to a series of linked nucleotide residues, which oligonucleotide has a sufficient number of nucleotide bases to be used in a PCR reaction. A short oligonucleotide sequence may be based on, or designed from, a genomic or cDNA sequence and is used to amplify, confirm, or reveal the presence of an identical, similar or complementary DNA or RNA in a particular cell or tissue. Oligonucleotides comprise portions of a nucleic acid sequence having about 10 nt, 50 nt, or 100 nt in length, preferably about 15 nt to 30 nt in length. In one embodiment, an oligonucleotide comprising a nucleic acid molecule less than 100 nt in length would further comprise at least 6 contiguous nucleotides of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or a complement thereof. Oligonucleotides may be chemically synthesized and may be used as probes.

In another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule that is a complement of the nucleotide sequence shown in any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161). In another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule that is a complement of the nucleotide sequence shown in any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or a portion of this nucleotide sequence. A nucleic acid molecule that is complementary to the nucleotide sequence shown in

is one that is sufficiently complementary to the nucleotide sequence shown in of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161) that it can hydrogen bond with little or no mismatches to the nucleotide sequence shown in of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161),, thereby forming a stable duplex.

5 As used herein, the term "complementary" refers to Watson-Crick or Hoogsteen base pairing between nucleotides units of a nucleic acid molecule, and the term "binding" means the physical or chemical interaction between two polypeptides or compounds or associated polypeptides or compounds or combinations thereof. Binding includes ionic, non-ionic, Von der Waals, hydrophobic interactions, etc. A physical interaction can be either direct or indirect.

10 Indirect interactions may be through or due to the effects of another polypeptide or compound. Direct binding refers to interactions that do not take place through, or due to, the effect of another polypeptide or compound, but instead are without other substantial chemical intermediates.

Moreover, the nucleic acid molecule of the invention can comprise only a portion of the
15 nucleic acid sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), *e.g.*, a fragment that can be used as a probe or primer, or a fragment encoding a biologically active portion of ORFX. Fragments provided herein are defined as sequences of at least 6 (contiguous) nucleic acids or at least 4 (contiguous) amino acids, a length sufficient to allow for specific hybridization in the case of nucleic acids or for specific recognition of an epitope in the case of amino acids, respectively,
20 and are at most some portion less than a full length sequence. Fragments may be derived from any contiguous portion of a nucleic acid or amino acid sequence of choice. Derivatives are nucleic acid sequences or amino acid sequences formed from the native compounds either directly or by modification or partial substitution. Analogs are nucleic acid sequences or amino acid sequences that have a structure similar to, but not identical to, the native compound but
25 differs from it in respect to certain components or side chains. Analogs may be synthetic or from a different evolutionary origin and may have a similar or opposite metabolic activity compared to wild type.

Derivatives and analogs may be full length or other than full length, if the derivative or analog contains a modified nucleic acid or amino acid, as described below. Derivatives or
30 analogs of the nucleic acids or proteins of the invention include, but are not limited to, molecules comprising regions that are substantially homologous to the nucleic acids or proteins of the

invention, in various embodiments, by at least about 70%, 80%, 85%, 90%, 95%, 98%, or even 99% identity (with a preferred identity of 80-99%) over a nucleic acid or amino acid sequence of identical size or when compared to an aligned sequence in which the alignment is done by a computer homology program known in the art, or whose encoding nucleic acid is capable of hybridizing to the complement of a sequence encoding the aforementioned proteins under stringent, moderately stringent, or low stringent conditions. See *e.g.* Ausubel, *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993, and below. An exemplary program is the Gap program (Wisconsin Sequence Analysis Package, Version 8 for UNIX, Genetics Computer Group, University Research Park, Madison, WI) using the default settings, which uses the algorithm of Smith and Waterman (Adv. Appl. Math., 1981, 2: 482-489, which is incorporated herein by reference in its entirety).

A "homologous nucleic acid sequence" or "homologous amino acid sequence," or variations thereof, refer to sequences characterized by a homology at the nucleotide level or amino acid level as discussed above. Homologous nucleotide sequences encode those sequences coding for isoforms of ORFX polypeptide. Isoforms can be expressed in different tissues of the same organism as a result of, for example, alternative splicing of RNA. Alternatively, isoforms can be encoded by different genes. In the present invention, homologous nucleotide sequences include nucleotide sequences encoding for a ORFX polypeptide of species other than humans, including, but not limited to, mammals, and thus can include, *e.g.*, mouse, rat, rabbit, dog, cat, cow, horse, and other organisms. Homologous nucleotide sequences also include, but are not limited to, naturally occurring allelic variations and mutations of the nucleotide sequences set forth herein. A homologous nucleotide sequence does not, however, include the nucleotide sequence encoding human ORFX protein. Homologous nucleic acid sequences include those nucleic acid sequences that encode conservative amino acid substitutions (see below) in any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161) as well as a polypeptide having ORFX activity. Biological activities of the ORFX proteins are described below. A homologous amino acid sequence does not encode the amino acid sequence of a human ORFX polypeptide.

The nucleotide sequence determined from the cloning of the human ORFX gene allows for the generation of probes and primers designed for use in identifying the cell types disclosed and/or cloning ORFX homologues in other cell types, *e.g.*, from other tissues, as well as ORFX homologues from other mammals. The probe/primer typically comprises a substantially purified

oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, 25, 50, 100, 150, 200, 250, 300, 350 or 400 or more consecutive sense strand nucleotide sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161); or an anti-sense strand nucleotide sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161);
5 or of a naturally occurring mutant of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161).

Probes based on the human ORFX nucleotide sequence can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In various embodiments, the probe further comprises a label group attached thereto, *e.g.*, the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be
10 used as a part of a diagnostic test kit for identifying cells or tissue which misexpress a ORFX protein, such as by measuring a level of a ORFX-encoding nucleic acid in a sample of cells from a subject *e.g.*, detecting ORFX mRNA levels or determining whether a genomic ORFX gene has been mutated or deleted.

"A polypeptide having a biologically active portion of ORFX" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. A nucleic acid fragment encoding a "biologically active portion of ORFX" can be prepared by isolating a portion of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), that encodes a polypeptide having a ORFX biological activity (biological activities of the ORFX
20 proteins are summarized in Table 1), expressing the encoded portion of ORFX protein (*e.g.*, by recombinant expression *in vitro*) and assessing the activity of the encoded portion of ORFX. For example, a nucleic acid fragment encoding a biologically active portion of ORFX can optionally include a domain as shown in Table 1, column 4.

25 ORFX variants

The invention further encompasses nucleic acid molecules that differ from the disclosed ORFX nucleotide sequences due to degeneracy of the genetic code. These nucleic acids thus encode the same ORFX protein as that encoded by the nucleotide sequence shown in SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161). In another embodiment, an isolated nucleic acid molecule of
30 the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161).

In addition to the human ORFX nucleotide sequence shown in any of SEQ ID NO:2n-1 (wherein $n = 1$ to 3161), it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of ORFX may exist within a population (e.g., the human population). Such genetic polymorphism in the ORFX gene may exist among individuals within a population due to natural allelic variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame encoding a ORFX protein, preferably a mammalian ORFX protein. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of the ORFX gene. Any and all such nucleotide variations and resulting amino acid polymorphisms in ORFX that are the result of natural allelic variation and that do not alter the functional activity of ORFX are intended to be within the scope of the invention.

Moreover, nucleic acid molecules encoding ORFX proteins from other species, and thus that have a nucleotide sequence that differs from the human sequence of any of SEQ ID NO:2n-1 (wherein $n = 1$ to 3161), are intended to be within the scope of the invention. Nucleic acid molecules corresponding to natural allelic variants and homologues of the ORFX cDNAs of the invention can be isolated based on their homology to the human ORFX nucleic acids disclosed herein using the human cDNAs, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions.

In another embodiment, an isolated nucleic acid molecule of the invention is at least 6 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of any of SEQ ID NO:2n-1 (wherein $n = 1$ to 3161). In another embodiment, the nucleic acid is at least 10, 25, 50, 100, 250, 500 or 750 nucleotides in length. In another embodiment, an isolated nucleic acid molecule of the invention hybridizes to the coding region. As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% homologous to each other typically remain hybridized to each other.

Homologs (i.e., nucleic acids encoding ORFX proteins derived from species other than human) or other related sequences (e.g., paralogs) can be obtained by low, moderate or high stringency hybridization with all or a portion of the particular human sequence as a probe using methods well known in the art for nucleic acid hybridization and cloning.

As used herein, the phrase "stringent hybridization conditions" refers to conditions under which a probe, primer or oligonucleotide will hybridize to its target sequence, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures than shorter sequences. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. Since the target sequences are generally present at excess, at T_m , 50% of the probes are occupied at equilibrium. Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes, primers or oligonucleotides (e.g., 10 nt to 50 nt) and at least about 60°C for longer probes, primers and oligonucleotides. Stringent conditions may also be achieved with the addition of destabilizing agents, such as formamide.

Stringent conditions are known to those skilled in the art and can be found in CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. Preferably, the conditions are such that sequences at least about 65%, 70%, 75%, 85%, 90%, 95%, 98%, or 99% homologous to each other typically remain hybridized to each other. A non-limiting example of stringent hybridization conditions is hybridization in a high salt buffer comprising 6X SSC, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA, and 500 mg/ml denatured salmon sperm DNA at 65°C. This hybridization is followed by one or more washes in 0.2X SSC, 0.01% BSA at 50°C. An isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of any of SEQ ID NO:2*n*-1 (wherein $n = 1$ to 3161) corresponds to a naturally occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (e.g., encodes a natural protein).

In a second embodiment, a nucleic acid sequence that is hybridizable to the nucleic acid molecule comprising the nucleotide sequence of any of SEQ ID NO:2*n*-1 (wherein $n = 1$ to 3161), or fragments, analogs or derivatives thereof, under conditions of moderate stringency is provided. A non-limiting example of moderate stringency hybridization conditions are

hybridization in 6X SSC, 5X Denhardt's solution, 0.5% SDS and 100 mg/ml denatured salmon sperm DNA at 55°C, followed by one or more washes in 1X SSC, 0.1% SDS at 37°C. Other conditions of moderate stringency that may be used are well known in the art. See, e.g., Ausubel *et al.* (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and
5 Kriegler, 1990, GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY.

In a third embodiment, a nucleic acid that is hybridizable to the nucleic acid molecule comprising the nucleotide sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or fragments, analogs or derivatives thereof, under conditions of low stringency, is provided. A
10 non-limiting example of low stringency hybridization conditions are hybridization in 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 mg/ml denatured salmon sperm DNA, 10% (wt/vol) dextran sulfate at 40°C, followed by one or more washes in 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS at 50°C. Other conditions of low stringency that may be used are well known in the art (e.g., as
15 employed for cross-species hybridizations). See, e.g., Ausubel *et al.* (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and Kriegler, 1990, GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY; Shilo and Weinberg, 1981, *Proc Natl Acad Sci USA* 78: 6789-6792.

Conservative mutations

In addition to naturally-occurring allelic variants of the ORFX sequence that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), thereby leading to changes in the amino acid sequence of the encoded ORFX protein, without altering the functional ability of the ORFX protein. For example, nucleotide substitutions
20 leading to amino acid substitutions at "non-essential" amino acid residues can be made in the sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161). A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of ORFX without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are conserved among the ORFX proteins of the present
30 invention, are predicted to be particularly unamenable to alteration.

Amino acid residues that are conserved among members of an ORFX family members are predicted to be less amenable to alteration. For example, an ORFX protein according to the present invention can contain at least one domain (*e.g.*, as shown in Table 1) that is a typically conserved region in an ORFX family member. As such, these conserved domains are not likely to be amenable to mutation. Other amino acid residues, however, (*e.g.*, those that are not conserved or only semi-conserved among members of the ORFX family) may not be as essential for activity and thus are more likely to be amenable to alteration.

Another aspect of the invention pertains to nucleic acid molecules encoding ORFX proteins that contain changes in amino acid residues that are not essential for activity. Such ORFX proteins differ in amino acid sequence from any of any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161), yet retain biological activity. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 75% homologous to the amino acid sequence of any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161). Preferably, the protein encoded by the nucleic acid is at least about 80% homologous to any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161), more preferably at least about 90%, 95%, 98%, and most preferably at least about 99% homologous to SEQ ID NO:2.

An isolated nucleic acid molecule encoding a ORFX protein homologous to the protein of any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161) can be created by introducing one or more nucleotide substitutions, additions or deletions into the corresponding nucleotide sequence, *i.e.* SEQ ID NO:2*n*-1 for the corresponding *n*, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein.

Mutations can be introduced into SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161) by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline,

phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in ORFX is replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a ORFX coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for ORFX biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), the encoded protein can be expressed by any recombinant technology known in the art and the activity of the protein can be determined.

In one embodiment, a mutant ORFX protein can be assayed for (1) the ability to form protein:protein interactions with other ORFX proteins, other cell-surface proteins, or biologically active portions thereof, (2) complex formation between a mutant ORFX protein and a ORFX receptor; (3) the ability of a mutant ORFX protein to bind to an intracellular target protein or biologically active portion thereof; (*e.g.*, avidin proteins); (4) the ability to bind BRA protein; or (5) the ability to specifically bind an anti-ORFX protein antibody.

Antisense

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire ORFX coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a ORFX protein of any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161) or antisense nucleic acids complementary to a ORFX nucleic acid sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161) are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding ORFX. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid

residues (*e.g.*, the protein coding region of a human ORFX that corresponds to any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161)). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding ORFX. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding ORFX disclosed herein (*e.g.*, SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161)), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of ORFX mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of ORFX mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of ORFX mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil,

3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a ORFX protein to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α -anomeric nucleic acid molecule. An α -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β -units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids Res* 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res* 15: 6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.* (1987) *FEBS Lett* 215: 327-330).

Ribozymes and PNA moieties

Such modifications include, by way of nonlimiting example, modified bases, and nucleic acids whose sugar phosphate backbones are modified or derivatized. These modifications are

carried out at least in part to enhance the chemical stability of the modified nucleic acid, such that they may be used, for example, as antisense binding nucleic acids in therapeutic applications in a subject.

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme.

5 Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave ORFX mRNA transcripts to thereby inhibit translation of ORFX mRNA. A ribozyme having specificity for a ORFX-encoding
10 nucleic acid can be designed based upon the nucleotide sequence of a ORFX DNA disclosed herein (*i.e.*, SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161)). For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a ORFX-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Pat. No. 4,987,071; and Cech *et al.* U.S. Pat. No. 5,116,742.

15 Alternatively, ORFX mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel *et al.*, (1993) *Science* 261:1411-1418.

Alternatively, ORFX gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the ORFX (*e.g.*, the ORFX promoter and/or
20 enhancers) to form triple helical structures that prevent transcription of the ORFX gene in target cells. See generally, Helene. (1991) *Anticancer Drug Des.* 6: 569-84; Helene. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14: 807-15.

In various embodiments, the nucleic acids of ORFX can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of
25 the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup *et al.* (1996) *Bioorg Med Chem* 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has
30 been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide

synthesis protocols as described in Hyrup *et al.* (1996) above; Perry-O'Keefe *et al.* (1996) *PNAS* 93: 14670-675.

PNAs of ORFX can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs of ORFX can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup *et al.* (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of ORFX can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of ORFX can be generated that may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn *et al.* (1996) *Nucl Acids Res* 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag *et al.* (1989) *Nucl Acid Res* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.* (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen *et al.* (1975) *Bioorg Med Chem Lett* 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.*, 1989, *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-6556; Lemaitre *et al.*, 1987, *Proc. Natl. Acad. Sci.* 84:648-652; PCT Publication No. W088/09810) or

the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, *e.g.*, Krol *et al.*, 1988, *BioTechniques* 6:958-976) or intercalating agents. (See, *e.g.*, Zon, 1988, *Pharm. Res.* 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

ORFX polypeptides

The novel protein of the invention includes the ORFX-like protein whose sequence is provided in any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161). The invention also includes a mutant or variant protein any of whose residues may be changed from the corresponding residue shown in FIG. 1 while still encoding a protein that maintains its ORFX-like activities and physiological functions, or a functional fragment thereof. For example, the invention includes the polypeptides encoded by the variant ORFX nucleic acids described above. In the mutant or variant protein, up to 20% or more of the residues may be so changed.

In general, an ORFX-like variant that preserves ORFX-like function includes any variant in which residues at a particular position in the sequence have been substituted by other amino acids, and further include the possibility of inserting an additional residue or residues between two residues of the parent protein as well as the possibility of deleting one or more residues from the parent sequence. Any amino acid substitution, insertion, or deletion is encompassed by the invention. In favorable circumstances, the substitution is a conservative substitution as defined above. Furthermore, without limiting the scope of the invention, positions of any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161) may be substituted such that a mutant or variant protein may include one or more substitutions

The invention also includes isolated ORFX proteins, and biologically active portions thereof, or derivatives, fragments, analogs or homologs thereof. Also provided are polypeptide fragments suitable for use as immunogens to raise anti-ORFX antibodies. In one embodiment, native ORFX proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, ORFX proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a ORFX

protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the ORFX protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of ORFX protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of ORFX protein having less than about 30% (by dry weight) of non-ORFX protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-ORFX protein, still more preferably less than about 10% of non-ORFX protein, and most preferably less than about 5% non-ORFX protein. When the ORFX protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of ORFX protein in which the protein is separated from chemical precursors or other chemicals that are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of ORFX protein having less than about 30% (by dry weight) of chemical precursors or non-ORFX chemicals, more preferably less than about 20% chemical precursors or non-ORFX chemicals, still more preferably less than about 10% chemical precursors or non-ORFX chemicals, and most preferably less than about 5% chemical precursors or non-ORFX chemicals.

Biologically active portions of a ORFX protein include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequence of the ORFX protein, *e.g.*, the amino acid sequence shown in SEQ ID NO:2 that include fewer amino acids than the full length ORFX proteins, and exhibit at least one activity of a ORFX protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the ORFX protein. A biologically active portion of a ORFX protein can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acids in length.

A biologically active portion of a ORFX protein of the present invention may contain at least one of the above-identified domains conserved between the FGF family of proteins.

Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native ORFX protein.

In an embodiment, the ORFX protein has an amino acid sequence shown in any of SEQ ID NO:2n (wherein n = 1 to 3161). In other embodiments, the ORFX protein is substantially homologous to any of SEQ ID NO:2n (wherein n = 1 to 3161) and retains the functional activity of the protein of any of SEQ ID NO:2n (wherein n = 1 to 3161), yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail below.

Accordingly, in another embodiment, the ORFX protein is a protein that comprises an amino acid sequence at least about 45% homologous, and more preferably about 55, 65, 70, 75, 80, 85, 90, 95, 98 or even 99% homologous to the amino acid sequence of any of SEQ ID NO:2n (wherein n = 1 to 3161) and retains the functional activity of the ORFX proteins of the corresponding polypeptide having the sequence of SEQ ID NO:2n (wherein n = 1 to 3161).

Determining homology between two or more sequences

To determine the percent homology of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in either of the sequences being compared for optimal alignment between the sequences). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are homologous at that position (*i.e.*, as used herein amino acid or nucleic acid "homology" is equivalent to amino acid or nucleic acid "identity").

The nucleic acid sequence homology may be determined as the degree of identity between two sequences. The homology may be determined using computer programs known in the art, such as GAP software provided in the GCG program package. See, *Needleman and Wunsch* 1970 *J Mol Biol* 48: 443-453. Using GCG GAP software with the following settings for nucleic acid sequence comparison: GAP creation penalty of 5.0 and GAP extension penalty of 0.3, the coding region of the analogous nucleic acid sequences referred to above exhibits a

11/12/2004

degree of identity preferably of at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99%, with the CDS (encoding) part of the DNA sequence shown in SEQ ID NO:2n-1 (wherein n = 1 to 3161).

The term "sequence identity" refers to the degree to which two polynucleotide or polypeptide sequences are identical on a residue-by-residue basis over a particular region of comparison. The term "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over that region of comparison, determining the number of positions at which the identical nucleic acid base (*e.g.*, A, T, C, G, U, or I, in the case of nucleic acids) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the region of comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of sequence identity. The term "substantial identity" as used herein denotes a characteristic of a polynucleotide sequence, wherein the polynucleotide comprises a sequence that has at least 80 percent sequence identity, preferably at least 85 percent identity and often 90 to 95 percent sequence identity, more usually at least 99 percent sequence identity as compared to a reference sequence over a comparison region. The term "percentage of positive residues" is calculated by comparing two optimally aligned sequences over that region of comparison, determining the number of positions at which the identical and conservative amino acid substitutions, as defined above, occur in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the region of comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of positive residues.

Chimeric and fusion proteins

The invention also provides ORFX chimeric or fusion proteins. As used herein, a ORFX "chimeric protein" or "fusion protein" includes a ORFX polypeptide operatively linked to a non-ORFX polypeptide. A "ORFX polypeptide" refers to a polypeptide having an amino acid sequence corresponding to ORFX, whereas a "non-ORFX polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein that is not substantially homologous to the ORFX protein, *e.g.*, a protein that is different from the ORFX protein and that is derived from the same or a different organism. Within a ORFX fusion protein the ORFX polypeptide can correspond to all or a portion of a ORFX protein. In one embodiment, a ORFX fusion protein comprises at least one biologically active portion of a ORFX protein. In another embodiment, a ORFX fusion protein comprises at least two biologically active portions of a

ORFX protein. Within the fusion protein, the term "operatively linked" is intended to indicate that the ORFX polypeptide and the non-ORFX polypeptide are fused in-frame to each other. The non-ORFX polypeptide can be fused to the N-terminus or C-terminus of the ORFX polypeptide.

For example, in one embodiment a ORFX fusion protein comprises a ORFX polypeptide operably linked to the extracellular domain of a second protein. Such fusion proteins can be further utilized in screening assays for compounds that modulate ORFX activity (such assays are described in detail below).

In another embodiment, the fusion protein is a GST-ORFX fusion protein in which the ORFX sequences are fused to the C-terminus of the GST (i.e., glutathione S-transferase) sequences. Such fusion proteins can facilitate the purification of recombinant ORFX.

In yet another embodiment, the fusion protein is a ORFX protein containing a heterologous signal sequence at its N-terminus. For example, the native ORFX signal sequence can be removed and replaced with a signal sequence from another protein. In certain host cells (e.g., mammalian host cells), expression and/or secretion of ORFX can be increased through use of a heterologous signal sequence.

In another embodiment, the fusion protein is a ORFX-immunoglobulin fusion protein in which the ORFX sequences comprising one or more domains are fused to sequences derived from a member of the immunoglobulin protein family. The ORFX-immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ORFX ligand and a ORFX protein on the surface of a cell, to thereby suppress ORFX-mediated signal transduction *in vivo*. In one nonlimiting example, a contemplated ORFX ligand of the invention is an ORFX receptor. The ORFX-immunoglobulin fusion proteins can be used to modulate the bioavailability of a ORFX cognate ligand. Inhibition of the ORFX ligand/ORFX interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, as well as modulating (e.g., promoting or inhibiting) cell survival. Moreover, the ORFX-immunoglobulin fusion proteins of the invention can be used as immunogens to produce anti-ORFX antibodies in a subject, to purify ORFX ligands, and in screening assays to identify molecules that inhibit the interaction of ORFX with a ORFX ligand.

A ORFX chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, *e.g.*, by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A ORFX-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the ORFX protein.

ORFX agonists and antagonists

The present invention also pertains to variants of the ORFX proteins that function as either ORFX agonists (mimetics) or as ORFX antagonists. Variants of the ORFX protein can be generated by mutagenesis, *e.g.*, discrete point mutation or truncation of the ORFX protein. An agonist of the ORFX protein can retain substantially the same, or a subset of, the biological activities of the naturally occurring form of the ORFX protein. An antagonist of the ORFX protein can inhibit one or more of the activities of the naturally occurring form of the ORFX protein by, for example, competitively binding to a downstream or upstream member of a cellular signaling cascade which includes the ORFX protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the ORFX proteins.

Variants of the ORFX protein that function as either ORFX agonists (mimetics) or as ORFX antagonists can be identified by screening combinatorial libraries of mutants, *e.g.*, truncation mutants, of the ORFX protein for ORFX protein agonist or antagonist activity. In one

embodiment, a variegated library of ORFX variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of ORFX variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential ORFX sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g., for phage display) containing the set of ORFX sequences therein. There are a variety of methods which can be used to produce libraries of potential ORFX variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential ORFX sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, e.g., Narang (1983) *Tetrahedron* 39:3; Itakura *et al.* (1984) *Annu Rev Biochem* 53:323; Itakura *et al.* (1984) *Science* 198:1056; Ike *et al.* (1983) *Nucl Acid Res* 11:477.

Polypeptide libraries

In addition, libraries of fragments of the ORFX protein coding sequence can be used to generate a variegated population of ORFX fragments for screening and subsequent selection of variants of a ORFX protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a ORFX coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA that can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal and internal fragments of various sizes of the ORFX protein.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of ORFX proteins. The most widely used techniques, which are amenable to high throughput analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors,

transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique that enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify ORFX variants (Arkin and Yourvan (1992) PNAS 89:7811-7815; Delgrave *et al.* (1993) Protein Engineering 6:327-331).

Anti-ORFX Antibodies

The invention further encompasses antibodies and antibody fragments, such as F_{ab} or $(F_{ab})_2$, that bind immunospecifically to any of the proteins of the invention.

An isolated ORFX protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind ORFX using standard techniques for polyclonal and monoclonal antibody preparation. Full-length ORFX protein can be used. Alternatively, the invention provides antigenic peptide fragments of ORFX for use as immunogens. The antigenic peptide of ORFX comprises at least 4 amino acid residues of the amino acid sequence shown in any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161). The antigenic peptide encompasses an epitope of ORFX such that an antibody raised against the peptide forms a specific immune complex with ORFX. The antigenic peptide may comprise at least 6 aa residues, at least 8 aa residues, at least 10 aa residues, at least 15 aa residues, at least 20 aa residues, or at least 30 aa residues. In one embodiment of the invention, the antigenic peptide comprises a polypeptide comprising at least 6 contiguous amino acids of any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161).

In an embodiment of the invention, epitopes encompassed by the antigenic peptide are regions of ORFX that are located on the surface of the protein, *e.g.*, hydrophilic regions. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, *e.g.*, Hopp and Woods, 1981, Proc. Nat. Acad. Sci. USA 78: 3824-3828; Kyte and Doolittle 1982, J. Mol. Biol. 157: 105-142, each incorporated herein by reference in their entirety.

As disclosed herein, an ORFX protein sequence of any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161), or derivatives, fragments, analogs or homologs thereof, may be utilized as immunogens in the generation of antibodies that immunospecifically-bind these protein components. The term "antibody" as used herein refers to immunoglobulin molecules and

immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen, such as ORFX. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab} and $F_{(ab)_2}$ fragments, and an F_{ab} expression library. In a specific embodiment, antibodies to human ORFX proteins are disclosed. Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies to a ORFX protein sequence of any of SEQ ID NO:2n (wherein n = 1 to 3161) or derivative, fragment, analog or homolog thereof. Some of these proteins are discussed below.

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by injection with the native protein, or a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, recombinantly expressed ORFX protein or a chemically synthesized ORFX polypeptide. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), human adjuvants such as *Bacille Calmette-Guerin* and *Corynebacterium parvum*, or similar immunostimulatory agents. If desired, the antibody molecules directed against ORFX can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as protein A chromatography to obtain the IgG fraction.

The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of ORFX. A monoclonal antibody composition thus typically displays a single binding affinity for a particular ORFX protein with which it immunoreacts. For preparation of monoclonal antibodies directed towards a particular ORFX protein, or derivatives, fragments, analogs or homologs thereof, any technique that provides for the production of antibody molecules by continuous cell line culture may be utilized. Such techniques include, but are not limited to, the hybridoma technique (see Kohler & Milstein, 1975 *Nature* 256: 495-497); the trioma technique; the human B-cell hybridoma technique (see Kozbor, *et al.*, 1983 *Immunol Today* 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, *et al.*, 1985 In: MONOCLONAL ANTIBODIES AND

CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, *et al.*, 1983. *Proc Natl Acad Sci USA* 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus *in vitro* (see Cole, *et al.*, 1985 In: MONOCLONAL ANTIBODIES AND
5 CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Each of the above citations are incorporated herein by reference in their entirety

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to a ORFX protein (see *e.g.*, U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see *e.g.*, Huse, *et al.*,
10 1989 *Science* 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a ORFX protein or derivatives, fragments, analogs or homologs thereof. Non-human antibodies can be "humanized" by techniques well known in the art. See *e.g.*, U.S. Patent No. 5,225,539. Each of the above citations are incorporated herein by reference. Antibody fragments that contain the idiotypes to a ORFX protein may be produced by
15 techniques known in the art including, but not limited to: (i) an F_{(ab)₂} fragment produced by pepsin digestion of an antibody molecule; (ii) an F_{ab} fragment generated by reducing the disulfide bridges of an F_{(ab)₂} fragment; (iii) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_v fragments.

Additionally, recombinant anti-ORFX antibodies, such as chimeric and humanized
20 monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in PCT International Application No. PCT/US86/02269; European Patent Application No. 184,187; European Patent
25 Application No. 171,496; European Patent Application No. 173,494; PCT International Publication No. WO 86/01533; U.S. Pat. No. 4,816,567; European Patent Application No. 125,023; Better *et al.* (1988) *Science* 240:1041-1043; Liu *et al.* (1987) *PNAS* 84:3439-3443; Liu *et al.* (1987) *J Immunol.* 139:3521-3526; Sun *et al.* (1987) *PNAS* 84:214-218; Nishimura *et al.* (1987) *Cancer Res* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; Shaw *et al.* (1988), *J. Natl Cancer Inst* 80:1553-1559; Morrison (1985) *Science* 229:1202-1207; Oi *et al.* (1986) *BioTechniques* 4:214; U.S. Pat. No. 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525;

Verhoeyan *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J Immunol* 141:4053-4060. Each of the above citations are incorporated herein by reference.

In one embodiment, methods for the screening of antibodies that possess the desired specificity include, but are not limited to, enzyme-linked immunosorbent assay (ELISA) and other immunologically-mediated techniques known within the art. In a specific embodiment, selection of antibodies that are specific to a particular domain of a ORFX protein is facilitated by generation of hybridomas that bind to the fragment of a ORFX protein possessing such a domain. Antibodies that are specific for one or more domains within a ORFX protein, *e.g.*, the domain spanning the first fifty amino-terminal residues specific to ORFX when compared to FGF-9, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

Anti-ORFX antibodies may be used in methods known within the art relating to the localization and/or quantitation of a ORFX protein (*e.g.*, for use in measuring levels of the ORFX protein within appropriate physiological samples, for use in diagnostic methods, for use in imaging the protein, and the like). In a given embodiment, antibodies for ORFX proteins, or derivatives, fragments, analogs or homologs thereof, that contain the antibody derived binding domain, are utilized as pharmacologically-active compounds [hereinafter "Therapeutics"].

An anti-ORFX antibody (*e.g.*, monoclonal antibody) can be used to isolate ORFX by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-ORFX antibody can facilitate the purification of natural ORFX from cells and of recombinantly produced ORFX expressed in host cells. Moreover, an anti-ORFX antibody can be used to detect ORFX protein (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the ORFX protein. Anti-ORFX antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or

phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I , ^{131}I , ^{35}S or ^3H .

ORFX Recombinant Vectors and Host Cells

5 Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding ORFX protein, or derivatives, fragments, analogs or homologs thereof. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can
10 be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are
15 replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the
20 invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the
25 recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, that is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner that allows for expression of the nucleotide sequence (e.g., in an *in vitro* transcription/translation
30 system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements

(e.g., polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Regulatory sequences include those that direct constitutive expression of a nucleotide sequence in many types of host cell and those that direct expression of the nucleotide sequence only in certain host cells (e.g., tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (e.g., ORFX proteins, mutant forms of ORFX, fusion proteins, etc.).

The recombinant expression vectors of the invention can be designed for expression of ORFX in prokaryotic or eukaryotic cells. For example, ORFX can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: (1) to increase expression of recombinant protein; (2) to increase the solubility of the recombinant protein; and (3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, Mass.) and pRIT5 (Pharmacia, Piscataway, N.J.) that fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amrann *et al.*, (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.*, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 60-89).

5 One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein. See, Gottesman, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 119-128. Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons
10 for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the ORFX expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari, *et al.*,
15 (1987) *EMBO J* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz *et al.*, (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, Calif.), and picZ (Invitrogen Corp, San Diego, Calif.).

Alternatively, ORFX can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*,
20 SF9 cells) include the pAc series (Smith *et al.* (1983) *Mol Cell Biol* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed (1987) *Nature* 329:840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J*
25 6: 187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells. See, *e.g.*, Chapters 16 and 17 of Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory,
30 Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.* (1987) *Genes Dev* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv Immunol* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J* 8:729-733) and immunoglobulins (Banerji *et al.* (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (*e.g.*, the neurofilament promoter; Byrne and Ruddle (1989) *PNAS* 86:5473-5477), pancreas-specific promoters (Edlund *et al.* (1985) *Science* 230:912-916), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Pat. No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, *e.g.*, the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the α -fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner that allows for expression (by transcription of the DNA molecule) of an RNA molecule that is antisense to ORFX mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen that direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen that direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub *et al.*, "Antisense RNA as a molecular tool for genetic analysis," Reviews--Trends in Genetics, Vol. 1(1) 1986.

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant

host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, ORFX protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (MOLECULAR CLONING: A LABORATORY MANUAL, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Various selectable markers include those that confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding ORFX or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) ORFX protein. Accordingly, the invention further provides methods for producing ORFX protein using the host cells of the invention. In one embodiment,

the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding ORFX has been introduced) in a suitable medium such that ORFX protein is produced. In another embodiment, the method further comprises isolating ORFX from the medium or the host cell.

5 Transgenic animals

The host cells of the invention can also be used to produce nonhuman transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which ORFX-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous ORFX sequences
10 have been introduced into their genome or homologous recombinant animals in which endogenous ORFX sequences have been altered. Such animals are useful for studying the function and/or activity of ORFX and for identifying and/or evaluating modulators of ORFX activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal
15 includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA that is integrated into the genome of a cell from which a transgenic animal develops and that remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous
20 recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous ORFX gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing ORFX-encoding
25 nucleic acid into the male pronuclei of a fertilized oocyte, *e.g.*, by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The human ORFX DNA sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161) can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homologue of the human ORFX gene, such as a mouse ORFX gene, can be isolated based on hybridization to
30 the human ORFX cDNA (described further above) and used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of

expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to the ORFX transgene to direct expression of ORFX protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Pat. Nos. 4,736,866; 4,870,009; and 4,873,191; and Hogan 1986, In: MANIPULATING THE MOUSE EMBRYO, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the ORFX transgene in its genome and/or expression of ORFX mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding ORFX can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which contains at least a portion of a ORFX gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the ORFX gene. The ORFX gene can be a human gene (*e.g.*, SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161)), but more preferably, is a non-human homologue of a human ORFX gene. For example, a mouse homologue of human ORFX gene of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161) can be used to construct a homologous recombination vector suitable for altering an endogenous ORFX gene in the mouse genome. In one embodiment, the vector is designed such that, upon homologous recombination, the endogenous ORFX gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector).

Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous ORFX gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous ORFX protein). In the homologous recombination vector, the altered portion of the ORFX gene is flanked at its 5' and 3' ends by additional nucleic acid of the ORFX gene to allow for homologous recombination to occur between the exogenous ORFX gene carried by the vector and an endogenous ORFX gene in an embryonic stem cell. The additional flanking ORFX nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector. See *e.g.*, Thomas *et al.* (1987) *Cell* 51:503 for a description of

homologous recombination vectors. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced ORFX gene has homologously recombined with the endogenous ORFX gene are selected (see e.g., Li *et al.* (1992) *Cell* 69:915).

The selected cells are then injected into a blastocyst of an animal (e.g., a mouse) to form aggregation chimeras. See e.g., Bradley 1987, In: TERATOCARCINOMAS AND EMBRYONIC STEM CELLS: A PRACTICAL APPROACH, Robertson, ed. IRL, Oxford, pp. 113-152. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley (1991) *Curr Opin Biotechnol* 2:823-829; PCT International Publication Nos.: WO 90/11354; WO 91/01140; WO 92/0968; and WO 93/04169.

In another embodiment, transgenic non-humans animals can be produced that contain selected systems that allow for regulated expression of the transgene. One example of such a system is the cre/loxP recombinase system of bacteriophage P1. For a description of the cre/loxP recombinase system, see, e.g., Lakso *et al.* (1992) *PNAS* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman *et al.* (1991) *Science* 251:1351-1355. If a cre/loxP recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the Cre recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmot *et al.* (1997) *Nature* 385:810-813. In brief, a cell, e.g., a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G₀ phase. The quiescent cell can then be fused, e.g., through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of

this female foster animal will be a clone of the animal from which the cell, *e.g.*, the somatic cell, is isolated.

Pharmaceutical Compositions

The ORFX nucleic acid molecules, ORFX proteins, and anti-ORFX antibodies (also referred to herein as "active compounds") of the invention, and derivatives, fragments, analogs and homologs thereof, can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein, "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Suitable carriers are described in the most recent edition of Remington's Pharmaceutical Sciences, a standard reference text in the field, which is incorporated herein by reference. Preferred examples of such carriers or diluents include, but are not limited to, water, saline, finger's solutions, dextrose solution, and 5% human serum albumin. Liposomes and non-aqueous vehicles such as fixed oils may also be used. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates, and agents for the adjustment of tonicity such as sodium chloride or dextrose. The pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, N.J.) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringeability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (e.g., a ORFX protein or anti-ORFX antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle that contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, methods of preparation are vacuum drying and freeze-drying that yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use

as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Pat. No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by any of a number of routes, *e.g.*, as described in U.S. Patent Nos. 5,703,055. Delivery can thus also include, *e.g.*, intravenous injection, local administration (see U.S. Pat. No. 5,328,470) or stereotactic injection (see *e.g.*, Chen *et al.* (1994) *PNAS* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the pharmaceutical preparation can include one or more cells that produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

Additional Uses and Methods of the Invention

The nucleic acid molecules, proteins, protein homologues, and antibodies described herein can be used in one or more of the following methods: (a) screening assays; (b) detection assays (*e.g.*, chromosomal mapping, cell and tissue typing, forensic biology), (c) predictive medicine (*e.g.*, diagnostic assays, prognostic assays, monitoring clinical trials, and pharmacogenomics); and (d) methods of treatment (*e.g.*, therapeutic and prophylactic).

The isolated nucleic acid molecules of the invention can be used to express ORFX protein (*e.g.*, via a recombinant expression vector in a host cell in gene therapy applications), to detect ORFX mRNA (*e.g.*, in a biological sample) or a genetic lesion in a ORFX gene, and to modulate ORFX activity, as described further below. In addition, the ORFX proteins can be used to screen drugs or compounds that modulate the ORFX activity or expression as well as to treat disorders characterized by insufficient or excessive production of ORFX protein, for

example proliferative or differentiative disorders, or production of ORFX protein forms that have decreased or aberrant activity compared to ORFX wild type protein. In addition, the anti-ORFX antibodies of the invention can be used to detect and isolate ORFX proteins and modulate ORFX activity.

5 This invention further pertains to novel agents identified by the above described screening assays and uses thereof for treatments as described herein.

Screening Assays

 The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides,
10 peptidomimetics, small molecules or other drugs) that bind to ORFX proteins or have a stimulatory or inhibitory effect on, for example, ORFX expression or ORFX activity.

 In one embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a ORFX protein or polypeptide or biologically active portion thereof. The test compounds of the present invention can be obtained
15 using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the "one-bead one-compound" library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are
20 applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam (1997) *Anticancer Drug Des* 12:145).

 Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc Natl Acad Sci U.S.A.* 90:6909; Erb *et al.* (1994) *Proc Natl Acad Sci U.S.A.* 91:11422; Zuckermann *et al.* (1994) *J Med Chem* 37:2678; Cho *et al.* (1993)
25 *Science* 261:1303; Carrell *et al.* (1994) *Angew Chem Int Ed Engl* 33:2059; Carell *et al.* (1994) *Angew Chem Int Ed Engl* 33:2061; and Gallop *et al.* (1994) *J Med Chem* 37:1233.

 Libraries of compounds may be presented in solution (*e.g.*, Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), on chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner U.S. Pat. No. 5,223,409), spores (Ladner USP '409),
30 plasmids (Cull *et al.* (1992) *Proc Natl Acad Sci USA* 89:1865-1869) or on phage (Scott and

Smith (1990) *Science* 249:386-390; Devlin (1990) *Science* 249:404-406; Cwirla *et al.* (1990) *Proc Natl Acad Sci U.S.A.* 87:6378-6382; Felici (1991) *J Mol Biol* 222:301-310; Ladner above.).

In one embodiment, an assay is a cell-based assay in which a cell which expresses a membrane-bound form of ORFX protein, or a biologically active portion thereof, on the cell surface is contacted with a test compound and the ability of the test compound to bind to a ORFX protein determined. The cell, for example, can of mammalian origin or a yeast cell. Determining the ability of the test compound to bind to the ORFX protein can be accomplished, for example, by coupling the test compound with a radioisotope or enzymatic label such that binding of the test compound to the ORFX protein or biologically active portion thereof can be determined by detecting the labeled compound in a complex. For example, test compounds can be labeled with ^{125}I , ^{35}S , ^{14}C , or ^3H , either directly or indirectly, and the radioisotope detected by direct counting of radioemission or by scintillation counting. Alternatively, test compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product. In one embodiment, the assay comprises contacting a cell which expresses a membrane-bound form of ORFX protein, or a biologically active portion thereof, on the cell surface with a known compound which binds ORFX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a ORFX protein, wherein determining the ability of the test compound to interact with a ORFX protein comprises determining the ability of the test compound to preferentially bind to ORFX or a biologically active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a membrane-bound form of ORFX protein, or a biologically active portion thereof, on the cell surface with a test compound and determining the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the ORFX protein or biologically active portion thereof. Determining the ability of the test compound to modulate the activity of ORFX or a biologically active portion thereof can be accomplished, for example, by determining the ability of the ORFX protein to bind to or interact with a ORFX target molecule. As used herein, a "target molecule" is a molecule with which a ORFX protein binds or interacts in nature, for example, a molecule on the surface of a cell which expresses a ORFX interacting protein, a molecule on the surface of a second cell, a molecule in the extracellular milieu, a molecule

associated with the internal surface of a cell membrane or a cytoplasmic molecule. A ORFX target molecule can be a non-ORFX molecule or a ORFX protein or polypeptide of the present invention. In one embodiment, a ORFX target molecule is a component of a signal transduction pathway that facilitates transduction of an extracellular signal (*e.g.*, a signal generated by binding of a compound to a membrane-bound ORFX molecule) through the cell membrane and into the cell. The target, for example, can be a second intercellular protein that has catalytic activity or a protein that facilitates the association of downstream signaling molecules with ORFX.

Determining the ability of the ORFX protein to bind to or interact with a ORFX target molecule can be accomplished by one of the methods described above for determining direct binding. In one embodiment, determining the ability of the ORFX protein to bind to or interact with a ORFX target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (*i.e.* intracellular Ca^{2+} , diacylglycerol, IP_3 , etc.), detecting catalytic/enzymatic activity of the target an appropriate substrate, detecting the induction of a reporter gene (comprising a ORFX-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, *e.g.*, luciferase), or detecting a cellular response, for example, cell survival, cellular differentiation, or cell proliferation.

In yet another embodiment, an assay of the present invention is a cell-free assay comprising contacting a ORFX protein or biologically active portion thereof with a test compound and determining the ability of the test compound to bind to the ORFX protein or biologically active portion thereof. Binding of the test compound to the ORFX protein can be determined either directly or indirectly as described above. In one embodiment, the assay comprises contacting the ORFX protein or biologically active portion thereof with a known compound which binds ORFX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a ORFX protein, wherein determining the ability of the test compound to interact with a ORFX protein comprises determining the ability of the test compound to preferentially bind to ORFX or biologically active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-free assay comprising contacting ORFX protein or biologically active portion thereof with a test compound and determining the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the ORFX protein or

biologically active portion thereof. Determining the ability of the test compound to modulate the activity of ORFX can be accomplished, for example, by determining the ability of the ORFX protein to bind to a ORFX target molecule by one of the methods described above for determining direct binding. In an alternative embodiment, determining the ability of the test compound to modulate the activity of ORFX can be accomplished by determining the ability of the ORFX protein further modulate a ORFX target molecule. For example, the catalytic/enzymatic activity of the target molecule on an appropriate substrate can be determined as previously described.

In yet another embodiment, the cell-free assay comprises contacting the ORFX protein or biologically active portion thereof with a known compound which binds ORFX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a ORFX protein, wherein determining the ability of the test compound to interact with a ORFX protein comprises determining the ability of the ORFX protein to preferentially bind to or modulate the activity of a ORFX target molecule.

The cell-free assays of the present invention are amenable to use of both the soluble form or the membrane-bound form of ORFX. In the case of cell-free assays comprising the membrane-bound form of ORFX, it may be desirable to utilize a solubilizing agent such that the membrane-bound form of ORFX is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton[®] X-100, Triton[®] X-114, Thesit[®], Isotridecypoly(ethylene glycol ether)_n, N-dodecyl--N,N-dimethyl-3-ammonio-1-propane sulfonate, 3-(3-cholamidopropyl)dimethylamminiol-1-propane sulfonate (CHAPS), or 3-(3-cholamidopropyl)dimethylamminiol-2-hydroxy-1-propane sulfonate (CHAPSO).

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either ORFX or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to ORFX, or interaction of ORFX with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided that adds a domain that allows one or both of the proteins to be bound to a matrix. For

example, GST-ORFX fusion proteins or GST-target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtiter plates, that are then combined with the test compound or the test compound and either the non-adsorbed target protein or ORFX protein, and the mixture is incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads; complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of ORFX binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either ORFX or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated ORFX or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals, Rockford, Ill.), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with ORFX or target molecules, but which do not interfere with binding of the ORFX protein to its target molecule, can be derivatized to the wells of the plate, and unbound target or ORFX trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the ORFX or target molecule, as well as enzyme-linked assays that rely on detecting an enzymatic activity associated with the ORFX or target molecule.

In another embodiment, modulators of ORFX expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of ORFX mRNA or protein in the cell is determined. The level of expression of ORFX mRNA or protein in the presence of the candidate compound is compared to the level of expression of ORFX mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of ORFX expression based on this comparison. For example, when expression of ORFX mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of ORFX mRNA or protein expression. Alternatively, when expression of ORFX mRNA or protein is less (statistically significantly less) in the presence of the candidate

compound than in its absence, the candidate compound is identified as an inhibitor of ORFX mRNA or protein expression. The level of ORFX mRNA or protein expression in the cells can be determined by methods described herein for detecting ORFX mRNA or protein.

In yet another aspect of the invention, the ORFX proteins can be used as "bait proteins" in a two-hybrid assay or three hybrid assay (see, *e.g.*, U.S. Pat. No. 5,283,317; Zervos *et al.* (1993) Cell 72:223-232; Madura *et al.* (1993) J Biol Chem 268:12046-12054; Bartel *et al.* (1993) Biotechniques 14:920-924; Iwabuchi *et al.* (1993) Oncogene 8:1693-1696; and Brent WO94/10300), to identify other proteins that bind to or interact with ORFX ("ORFX-binding proteins" or "ORFX-bp") and modulate ORFX activity. Such ORFX-binding proteins are also likely to be involved in the propagation of signals by the ORFX proteins as, for example, upstream or downstream elements of the ORFX pathway.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for ORFX is fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo*, forming a ORFX-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) that is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene that encodes the protein which interacts with ORFX.

This invention further pertains to novel agents identified by the above-described screening assays and uses thereof for treatments as described herein.

Detection Assays

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome; and, thus, locate gene regions associated with genetic disease; (ii) identify an individual from a

minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample.

The ORFX sequences of the present invention can also be used to identify individuals from minute biological samples. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. The sequences of the present invention are useful as additional DNA markers for RFLP ("restriction fragment length polymorphisms," described in U.S. Pat. No. 5,272,057).

Furthermore, the sequences of the present invention can be used to provide an alternative technique that determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the ORFX sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the present invention can be used to obtain such identification sequences from individuals and from tissue. The ORFX sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Much of the allelic variation is due to single nucleotide polymorphisms (SNPs), which include restriction fragment length polymorphisms (RFLPs).

Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences of SEQ ID NO:2 n -1 (wherein $n = 1$ to 3161), as described above, can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000 primers that each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

Predictive Medicine

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining ORFX protein and/or nucleic acid expression as well as ORFX activity, in the context of a biological sample (e.g., blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant ORFX expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with ORFX protein, nucleic acid expression or activity. For example, mutations in a ORFX gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with ORFX protein, nucleic acid expression or activity.

Another aspect of the invention provides methods for determining ORFX protein, nucleic acid expression or ORFX activity in an individual to thereby select appropriate therapeutic or prophylactic agents for that individual (referred to herein as "pharmacogenomics").

Pharmacogenomics allows for the selection of agents (e.g., drugs) for therapeutic or prophylactic treatment of an individual based on the genotype of the individual (e.g., the genotype of the individual examined to determine the ability of the individual to respond to a particular agent.)

Yet another aspect of the invention pertains to monitoring the influence of agents (e.g., drugs, compounds) on the expression or activity of ORFX in clinical trials.

Use of Partial ORFX Sequences in Forensic Biology

DNA-based identification techniques can also be used in forensic biology. Forensic biology is a scientific field employing genetic typing of biological evidence found at a crime scene as a means for positively identifying, for example, a perpetrator of a crime. To make such an identification, PCR technology can be used to amplify DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, or semen

found at a crime scene. The amplified sequence can then be compared to a standard, thereby allowing identification of the origin of the biological sample.

The sequences of the present invention can be used to provide polynucleotide reagents, *e.g.*, PCR primers, targeted to specific loci in the human genome, that can enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (*i.e.* another DNA sequence that is unique to a particular individual). As mentioned above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Sequences targeted to noncoding regions of SEQ ID NOs:___ are particularly appropriate for this use as greater numbers of polymorphisms occur in the noncoding regions, making it easier to differentiate individuals using this technique. Examples of polynucleotide reagents include the ORFX sequences or portions thereof, *e.g.*, fragments derived from the noncoding regions of one or more of SEQ ID NO:2*n*-1 (where *n* = 1 to 3161), having a length of at least 20 bases, preferably at least 30 bases.

The ORFX sequences described herein can further be used to provide polynucleotide reagents, *e.g.*, labeled or label-able probes that can be used, for example, in an *in situ* hybridization technique, to identify a specific tissue, *e.g.*, brain tissue, etc. This can be very useful in cases where a forensic pathologist is presented with a tissue of unknown origin. Panels of such ORFX probes can be used to identify tissue by species and/or by organ type.

In a similar fashion, these reagents, *e.g.*, ORFX primers or probes can be used to screen tissue culture for contamination (*i.e.* screen for the presence of a mixture of different types of cells in a culture).

Predictive Medicine

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining ORFX protein and/or nucleic acid expression as well as ORFX activity, in the context of a biological sample (*e.g.*, blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant ORFX expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with ORFX protein, nucleic

acid expression or activity. For example, mutations in a ORFX gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with ORFX protein, nucleic acid expression or activity.

5 Another aspect of the invention provides methods for determining ORFX protein, nucleic acid expression or ORFX activity in an individual to thereby select appropriate therapeutic or prophylactic agents for that individual (referred to herein as "pharmacogenomics").

Pharmacogenomics allows for the selection of agents (*e.g.*, drugs) for therapeutic or prophylactic treatment of an individual based on the genotype of the individual (*e.g.*, the genotype of the individual examined to determine the ability of the individual to respond to a particular agent.)

10 Yet another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of ORFX in clinical trials.

These and other agents are described in further detail in the following sections.

Diagnostic Assays

15 Other conditions in which proliferation of cells plays a role include tumors, restenosis, psoriasis, Dupuytren's contracture, diabetic complications, Kaposi's sarcoma and rheumatoid arthritis.

An ORFX polypeptide may be used to identify an interacting polypeptide a sample or tissue. The method comprises contacting the sample or tissue with ORFX, allowing formation of a complex between the ORFX polypeptide and the interacting polypeptide, and detecting the complex, if present.

20 The proteins of the invention may be used to stimulate production of antibodies specifically binding the proteins. Such antibodies may be used in immunodiagnostic procedures to detect the occurrence of the protein in a sample. The proteins of the invention may be used to stimulate cell growth and cell proliferation in conditions in which such growth would be favorable. An example would be to counteract toxic side effects of chemotherapeutic agents on, for example, hematopoiesis and platelet formation, linings of the gastrointestinal tract, and hair follicles. They may also be used to stimulate new cell growth in neurological disorders including, for example, Alzheimer's disease. Alternatively, antagonistic treatments may be administered in which an antibody specifically binding the ORFX-like proteins of the invention

25
30

would abrogate the specific growth-inducing effects of the proteins. Such antibodies may be useful, for example, in the treatment of proliferative disorders including various tumors and benign hyperplasias.

Polynucleotides or oligonucleotides corresponding to any one portion of the ORFX nucleic acids of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161) may be used to detect DNA containing a corresponding ORF gene, or detect the expression of a corresponding ORFX gene, or ORFX-like gene. For example, an ORFX nucleic acid expressed in a particular cell or tissue, as noted in Table 2, can be used to identify the presence of that particular cell type.

An exemplary method for detecting the presence or absence of ORFX in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting ORFX protein or nucleic acid (*e.g.*, mRNA, genomic DNA) that encodes ORFX protein such that the presence of ORFX is detected in the biological sample. An agent for detecting ORFX mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to ORFX mRNA or genomic DNA. The nucleic acid probe can be, for example, a full-length ORFX nucleic acid, such as the nucleic acid of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to ORFX mRNA or genomic DNA, as described above. Other suitable probes for use in the diagnostic assays of the invention are described herein.

An agent for detecting ORFX protein is an antibody capable of binding to ORFX protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or F(ab')₂) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect ORFX mRNA, protein, or genomic DNA in a biological sample *in vitro* as

well as *in vivo*. For example, *in vitro* techniques for detection of ORFX mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of ORFX protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. *In vitro* techniques for detection of ORFX genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of ORFX protein include introducing into a subject a labeled anti-ORFX antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting ORFX protein, mRNA, or genomic DNA, such that the presence of ORFX protein, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of ORFX protein, mRNA or genomic DNA in the control sample with the presence of ORFX protein, mRNA or genomic DNA in the test sample.

The invention also encompasses kits for detecting the presence of ORFX in a biological sample. For example, the kit can comprise: a labeled compound or agent capable of detecting ORFX protein or mRNA in a biological sample; means for determining the amount of ORFX in the sample; and means for comparing the amount of ORFX in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect ORFX protein or nucleic acid.

Prognostic Assays

The diagnostic methods described herein can furthermore be utilized to identify subjects having or at risk of developing a disease or disorder associated with aberrant ORFX expression or activity. For example, the assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with ORFX protein, nucleic acid expression or activity in, e.g., proliferative or differentiative disorders such as hyperplasias, tumors, restenosis, psoriasis, Dupuytren's

contracture, diabetic complications, or rheumatoid arthritis, etc.; and glia-associated disorders such as cerebral lesions, diabetic neuropathies, cerebral edema, senile dementia, Alzheimer's disease, etc. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a disease or disorder. Thus, the present invention provides a method for identifying a disease or disorder associated with aberrant ORFX expression or activity in which a test sample is obtained from a subject and ORFX protein or nucleic acid (*e.g.*, mRNA, genomic DNA) is detected, wherein the presence of ORFX protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant ORFX expression or activity. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (*e.g.*, serum), cell sample, or tissue.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant ORFX expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for a disorder, such as a proliferative disorder, differentiative disorder, glia-associated disorders, etc. Thus, the present invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant ORFX expression or activity in which a test sample is obtained and ORFX protein or nucleic acid is detected (*e.g.*, wherein the presence of ORFX protein or nucleic acid is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant ORFX expression or activity.)

The methods of the invention can also be used to detect genetic lesions in a ORFX gene, thereby determining if a subject with the lesioned gene is at risk for, or suffers from, a proliferative disorder, differentiative disorder, glia-associated disorder, etc. In various embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic lesion characterized by at least one of an alteration affecting the integrity of a gene encoding a ORFX-protein, or the mis-expression of the ORFX gene. For example, such genetic lesions can be detected by ascertaining the existence of at least one of (1) a deletion of one or more nucleotides from a ORFX gene; (2) an addition of one or more nucleotides to a ORFX gene; (3) a substitution of one or more nucleotides of a ORFX gene, (4) a chromosomal

rearrangement of a ORFX gene; (5) an alteration in the level of a messenger RNA transcript of a ORFX gene, (6) aberrant modification of a ORFX gene, such as of the methylation pattern of the genomic DNA, (7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a ORFX gene, (8) a non-wild type level of a ORFX-protein, (9) allelic loss of a ORFX gene, and (10) inappropriate post-translational modification of a ORFX-protein. As described herein, there are a large number of assay techniques known in the art which can be used for detecting lesions in a ORFX gene. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject. However, any biological sample containing nucleated cells may be used, including, for example, buccal mucosal cells.

In certain embodiments, detection of the lesion involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, e.g., U.S. Pat. Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, e.g., Landegran *et al.* (1988) *Science* 241:1077-1080; and Nakazawa *et al.* (1994) *PNAS* 91:360-364), the latter of which can be particularly useful for detecting point mutations in the ORFX-gene (see Abravaya *et al.* (1995) *Nucl Acids Res* 23:675-682). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers that specifically hybridize to a ORFX gene under conditions such that hybridization and amplification of the ORFX gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli *et al.*, 1990, *Proc Natl Acad Sci USA* 87:1874-1878), transcriptional amplification system (Kwoh, *et al.*, 1989, *Proc Natl Acad Sci USA* 86:1173-1177), Q-Beta Replicase (Lizardi *et al.*, 1988, *BioTechnology* 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in a ORFX gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Pat. No. 5,493,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in ORFX can be identified by hybridizing a sample and control nucleic acids, *e.g.*, DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotides probes (Cronin *et al.* (1996) *Human Mutation* 7: 244-255; Kozal *et al.* (1996) *Nature Medicine* 2: 753-759). For example, genetic mutations in ORFX can be identified in two dimensional arrays containing light-generated DNA probes as described in Cronin *et al.* above. Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the ORFX gene and detect mutations by comparing the sequence of the sample ORFX with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxim and Gilbert (1977) *PNAS* 74:560 or Sanger (1977) *PNAS* 74:5463. It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays (Naeve *et al.*, (1995) *Biotechniques* 19:448), including sequencing by mass spectrometry (see, *e.g.*, PCT International Publ. No. WO 94/16101; Cohen *et al.* (1996) *Adv Chromatogr* 36:127-162; and Griffin *et al.* (1993) *Appl Biochem Biotechnol* 38:147-159).

Other methods for detecting mutations in the ORFX gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA

heteroduplexes (Myers *et al.* (1985) *Science* 230:1242). In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labeled) RNA or DNA containing the wild-type ORFX sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent that cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digesting the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, for example, Cotton *et al* (1988) *Proc Natl Acad Sci USA* 85:4397; Saleeba *et al* (1992) *Methods Enzymol* 217:286-295. In an embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in ORFX cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu *et al.* (1994) *Carcinogenesis* 15:1657-1662). According to an exemplary embodiment, a probe based on a ORFX sequence, *e.g.*, a wild-type ORFX sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, for example, U.S. Pat. No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in ORFX genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita *et al.* (1989) *Proc Natl Acad Sci USA*: 86:2766, see also Cotton (1993) *Mutat Res* 285:125-144; Hayashi (1992) *Genet Anal Tech Appl* 9:73-79). Single-stranded DNA fragments of sample and control ORFX nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting

alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA, rather than DNA, in which the secondary structure is more sensitive to a change in sequence. In one embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility. See, e.g., Keen *et al.* (1991) *Trends Genet* 7:5.

In yet another embodiment the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE). See, e.g., Myers *et al.* (1985) *Nature* 313:495. When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA. See, e.g., Rosenbaum and Reissner (1987) *Biophys Chem* 265:12753.

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions that permit hybridization only if a perfect match is found. See, e.g., Saiki *et al.* (1986) *Nature* 324:163; Saiki *et al.* (1989) *Proc Natl Acad. Sci USA* 86:6230. Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology that depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs *et al.* (1989) *Nucleic Acids Res* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection. See, e.g., Gasparini *et al.* (1992) *Mol Cell Probes* 6:1. It is anticipated that in certain embodiments amplification may also be performed using Taq ligase

for amplification. See, *e.g.*, Barany (1991) *Proc Natl Acad Sci USA* 88:189. In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence, making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

5 The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, *e.g.*, in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a ORFX gene.

10 Furthermore, any cell type or tissue, preferably peripheral blood leukocytes, in which ORFX is expressed may be utilized in the prognostic assays described herein. However, any biological sample containing nucleated cells may be used, including, for example, buccal mucosal cells.

Pharmacogenomics

15 Agents, or modulators that have a stimulatory or inhibitory effect on ORFX activity (*e.g.*, ORFX gene expression), as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) disorders (*e.g.*, neurological, cancer-related or gestational disorders) associated with aberrant ORFX activity. In conjunction with such treatment, the pharmacogenomics (*i.e.*, the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) of the individual may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (*e.g.*, drugs) for prophylactic or therapeutic treatments based on a consideration of the individual's genotype. Such pharmacogenomics can further be used to determine appropriate dosages and therapeutic regimens. Accordingly, the activity of ORFX protein, expression of ORFX nucleic acid, or mutation content of ORFX genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

25 Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See *e.g.*,
30 Eichelbaum, 1996, *Clin Exp Pharmacol Physiol*, 23:983-985 and Linder, 1997, *Clin Chem*, 43:254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic

conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare defects or as polymorphisms. For example, glucose-6-phosphate dehydrogenase (G6PD) deficiency is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Thus, the activity of ORFX protein, expression of ORFX nucleic acid, or mutation content of ORFX genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual. In addition, pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding drug-metabolizing enzymes to the identification of an individual's drug responsiveness phenotype. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a ORFX modulator, such as a modulator identified by one of the exemplary screening assays described herein.

Monitoring Clinical Efficacy

Monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of ORFX (*e.g.*, the ability to modulate aberrant cell proliferation and/or differentiation) can be applied in basic drug screening and in clinical trials. For example, the effectiveness of an agent
5 determined by a screening assay as described herein to increase ORFX gene expression, protein levels, or upregulate ORFX activity, can be monitored in clinical trials of subjects exhibiting decreased ORFX gene expression, protein levels, or downregulated ORFX activity.

Alternatively, the effectiveness of an agent determined by a screening assay to decrease ORFX gene expression, protein levels, or downregulate ORFX activity, can be monitored in clinical
10 trials of subjects exhibiting increased ORFX gene expression, protein levels, or upregulated ORFX activity. In such clinical trials, the expression or activity of ORFX and, preferably, other genes that have been implicated in, for example, a proliferative or neurological disorder, can be used as a "read out" or marker of the responsiveness of a particular cell.

For example, genes, including ORFX, that are modulated in cells by treatment with an
15 agent (*e.g.*, compound, drug or small molecule) that modulates ORFX activity (*e.g.*, identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on cellular proliferation disorders, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of ORFX and other genes implicated in the disorder. The levels of gene expression (*i.e.*, a gene expression pattern) can be quantified by
20 Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of ORFX or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during, treatment of the individual with the
25 agent.

In one embodiment, the invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, protein, peptide, nucleic acid, peptidomimetic, small molecule, or other drug candidate identified by the screening assays described herein) comprising the steps of (i) obtaining a pre-administration sample from a
30 subject prior to administration of the agent; (ii) detecting the level of expression of a ORFX protein, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more

post-administration samples from the subject; (iv) detecting the level of expression or activity of the ORFX protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the ORFX protein, mRNA, or genomic DNA in the pre-administration sample with the ORFX protein, mRNA, or genomic DNA in the post administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of ORFX to higher levels than detected, *i.e.*, to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of ORFX to lower levels than detected, *i.e.*, to decrease the effectiveness of the agent.

Methods of Treatment

The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant ORFX expression or activity.

Diseases and disorders that are characterized by increased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that antagonize (*i.e.*, reduce or inhibit) activity. Therapeutics that antagonize activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to, (i) a ORFX polypeptide, or analogs, derivatives, fragments or homologs thereof; (ii) antibodies to a ORFX peptide; (iii) nucleic acids encoding a ORFX peptide; (iv) administration of antisense nucleic acid and nucleic acids that are "dysfunctional" (*i.e.*, due to a heterologous insertion within the coding sequences of coding sequences to a ORFX peptide) that are utilized to "knockout" endogenous function of a ORFX peptide by homologous recombination (see, *e.g.*, Capecchi, 1989, *Science* 244: 1288-1292); or (v) modulators (*i.e.*, inhibitors, agonists and antagonists, including additional peptide mimetic of the invention or antibodies specific to a peptide of the invention) that alter the interaction between a ORFX peptide and its binding partner.

Diseases and disorders that are characterized by decreased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that increase (*i.e.*, are agonists to) activity. Therapeutics that upregulate activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized

include, but are not limited to, a ORFX peptide, or analogs, derivatives, fragments or homologs thereof; or an agonist that increases bioavailability.

Increased or decreased levels can be readily detected by quantifying peptide and/or RNA, by obtaining a patient tissue sample (*e.g.*, from biopsy tissue) and assaying it *in vitro* for RNA or peptide levels, structure and/or activity of the expressed peptides (or mRNAs of a ORFX peptide). Methods that are well-known within the art include, but are not limited to, immunoassays (*e.g.*, by Western blot analysis, immunoprecipitation followed by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis, immunocytochemistry, etc.) and/or hybridization assays to detect expression of mRNAs (*e.g.*, Northern assays, dot blots, *in situ* hybridization, etc.).

In one aspect, the invention provides a method for preventing, in a subject, a disease or condition associated with an aberrant ORFX expression or activity, by administering to the subject an agent that modulates ORFX expression or at least one ORFX activity. Subjects at risk for a disease that is caused or contributed to by aberrant ORFX expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the ORFX aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of ORFX aberrancy, for example, a ORFX agonist or ORFX antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

Another aspect of the invention pertains to methods of modulating ORFX expression or activity for therapeutic purposes. The modulatory method of the invention involves contacting a cell with an agent that modulates one or more of the activities of ORFX protein activity associated with the cell. An agent that modulates ORFX protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring cognate ligand of a ORFX protein, a peptide, a ORFX peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more ORFX protein activity. Examples of such stimulatory agents include active ORFX protein and a nucleic acid molecule encoding ORFX that has been introduced into the cell. In another embodiment, the agent inhibits one or more ORFX protein activity. Examples of such inhibitory agents include antisense ORFX nucleic acid molecules and anti-ORFX antibodies. These modulatory methods can be performed *in vitro*

(e.g., by culturing the cell with the agent) or, alternatively, *in vivo* (e.g., by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of a ORFX protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (e.g., an agent identified by a screening assay described herein), or combination of agents that modulates (e.g., upregulates or downregulates) ORFX expression or activity. In another embodiment, the method involves administering a ORFX protein or nucleic acid molecule as therapy to compensate for reduced or aberrant ORFX expression or activity.

Determination of the Biological Effect of a Therapeutic

In various embodiments of the present invention, suitable *in vitro* or *in vivo* assays are utilized to determine the effect of a specific Therapeutic and whether its administration is indicated for treatment of the affected tissue.

In various specific embodiments, *in vitro* assays may be performed with representative cells of the type(s) involved in the patient's disorder, to determine if a given Therapeutic exerts the desired effect upon the cell type(s). Compounds for use in therapy may be tested in suitable animal model systems including, but not limited to rats, mice, chicken, cows, monkeys, rabbits, and the like, prior to testing in human subjects. Similarly, for *in vivo* testing, any of the animal model system known in the art may be used prior to administration to human subjects.

Malignancies

Some ORFX polypeptides are expressed in cancerous cells (*see, e.g.*, Tables 1 and 2). Accordingly, the corresponding ORF protein is involved in the regulation of cell proliferation. Accordingly, Therapeutics of the present invention may be useful in the therapeutic or prophylactic treatment of diseases or disorders that are associated with cell hyperproliferation and/or loss of control of cell proliferation (e.g., cancers, malignancies and tumors). For a review of such hyperproliferation disorders, *see e.g.*, Fishman, *et al.*, 1985. MEDICINE, 2nd ed., J.B. Lippincott Co., Philadelphia, PA.

Therapeutics of the present invention may be assayed by any method known within the art for efficacy in treating or preventing malignancies and related disorders. Such assays include,

but are not limited to, *in vitro* assays utilizing transformed cells or cells derived from the patient's tumor, as well as *in vivo* assays using animal models of cancer or malignancies. Potentially effective Therapeutics are those that, for example, inhibit the proliferation of tumor-derived or transformed cells in culture or cause a regression of tumors in animal models, in comparison to the controls.

In the practice of the present invention, once a malignancy or cancer has been shown to be amenable to treatment by modulating (*i.e.*, inhibiting, antagonizing or agonizing) activity, that cancer or malignancy may subsequently be treated or prevented by the administration of a Therapeutic that serves to modulate protein function.

10 Premalignant conditions

The Therapeutics of the present invention that are effective in the therapeutic or prophylactic treatment of cancer or malignancies may also be administered for the treatment of pre-malignant conditions and/or to prevent the progression of a pre-malignancy to a neoplastic or malignant state. Such prophylactic or therapeutic use is indicated in conditions known or suspected of preceding progression to neoplasia or cancer, in particular, where non-neoplastic cell growth consisting of hyperplasia, metaplasia or, most particularly, dysplasia has occurred. For a review of such abnormal cell growth see *e.g.*, Robbins & Angell, 1976. BASIC PATHOLOGY, 2nd ed., W.B. Saunders Co., Philadelphia, PA.

Hyperplasia is a form of controlled cell proliferation involving an increase in cell number in a tissue or organ, without significant alteration in its structure or function. For example, it has been demonstrated that endometrial hyperplasia often precedes endometrial cancer. Metaplasia is a form of controlled cell growth in which one type of mature or fully differentiated cell substitutes for another type of mature cell. Metaplasia may occur in epithelial or connective tissue cells. Dysplasia is generally considered a precursor of cancer, and is found mainly in the epithelia. Dysplasia is the most disorderly form of non-neoplastic cell growth, and involves a loss in individual cell uniformity and in the architectural orientation of cells. Dysplasia characteristically occurs where there exists chronic irritation or inflammation, and is often found in the cervix, respiratory passages, oral cavity, and gall bladder.

Alternatively, or in addition to the presence of abnormal cell growth characterized as hyperplasia, metaplasia, or dysplasia, the presence of one or more characteristics of a transformed or malignant phenotype displayed either *in vivo* or *in vitro* within a cell sample derived from a patient, is indicative of the desirability of prophylactic/therapeutic administration of a Therapeutic that possesses the ability to modulate activity of An aforementioned protein. Characteristics of a transformed phenotype include, but are not limited to: (i) morphological changes; (ii) looser substratum attachment; (iii) loss of cell-to-cell contact inhibition; (iv) loss of anchorage dependence; (v) protease release; (vi) increased sugar transport; (vii) decreased serum requirement; (viii) expression of fetal antigens, (ix) disappearance of the 250 kDal cell-surface protein, and the like. See *e.g.*, Richards, *et al.*, 1986. MOLECULAR PATHOLOGY, W.B. Saunders Co., Philadelphia, PA.

In a specific embodiment of the present invention, a patient that exhibits one or more of the following predisposing factors for malignancy is treated by administration of an effective amount of a Therapeutic: (i) a chromosomal translocation associated with a malignancy (*e.g.*, the Philadelphia chromosome (*bcr/abl*) for chronic myelogenous leukemia and t(14;18) for follicular lymphoma, etc.); (ii) familial polyposis or Gardner's syndrome (possible forerunners of colon cancer); (iii) monoclonal gammopathy of undetermined significance (a possible precursor of multiple myeloma) and (iv) a first degree kinship with persons having a cancer or pre-cancerous disease showing a Mendelian (genetic) inheritance pattern (*e.g.*, familial polyposis of the colon, Gardner's syndrome, hereditary exostosis, polyendocrine adenomatosis, Peutz-Jeghers syndrome, neurofibromatosis of Von Recklinghausen, medullary thyroid carcinoma with amyloid production and pheochromocytoma, retinoblastoma, carotid body tumor, cutaneous melanocarcinoma, intraocular melanocarcinoma, xeroderma pigmentosum, ataxia telangiectasia, Chediak-Higashi syndrome, albinism, Fanconi's aplastic anemia and Bloom's syndrome).

In another embodiment, a Therapeutic of the present invention is administered to a human patient to prevent the progression to breast, colon, lung, pancreatic, or uterine cancer, or melanoma or sarcoma.

Hyperproliferative and dysproliferative disorders

In one embodiment of the present invention, a Therapeutic is administered in the therapeutic or prophylactic treatment of hyperproliferative or benign dysproliferative disorders. The efficacy in treating or preventing hyperproliferative diseases or disorders of a Therapeutic of the present invention may be assayed by any method known within the art. Such assays include *in vitro* cell proliferation assays, *in vitro* or *in vivo* assays using animal models of hyperproliferative diseases or disorders, or the like. Potentially effective Therapeutics may, for example, promote cell proliferation in culture or cause growth or cell proliferation in animal models in comparison to controls.

Specific embodiments of the present invention are directed to the treatment or prevention of cirrhosis of the liver (a condition in which scarring has overtaken normal liver regeneration processes); treatment of keloid (hypertrophic scar) formation causing disfiguring of the skin in which the scarring process interferes with normal renewal; psoriasis (a common skin condition characterized by excessive proliferation of the skin and delay in proper cell fate determination); benign tumors; fibrocystic conditions and tissue hypertrophy (*e.g.*, benign prostatic hypertrophy).

Neurodegenerative disorders

Some ORFX proteins are found in cell types have been implicated in the deregulation of cellular maturation and apoptosis, which are both characteristic of neurodegenerative disease. Accordingly, Therapeutics of the invention, particularly but not limited to those that modulate (or supply) activity of an aforementioned protein, may be effective in treating or preventing neurodegenerative disease. Therapeutics of the present invention that modulate the activity of an aforementioned protein involved in neurodegenerative disorders can be assayed by any method known in the art for efficacy in treating or preventing such neurodegenerative diseases and disorders. Such assays include *in vitro* assays for regulated cell maturation or inhibition of apoptosis or *in vivo* assays using animal models of neurodegenerative diseases or disorders, or any of the assays described below. Potentially effective Therapeutics, for example but not by way of limitation, promote regulated cell maturation and prevent cell apoptosis in culture, or reduce neurodegeneration in animal models in comparison to controls.

Once a neurodegenerative disease or disorder has been shown to be amenable to treatment by modulation activity, that neurodegenerative disease or disorder can be treated or prevented by administration of a Therapeutic that modulates activity. Such diseases include all degenerative disorders involved with aging, especially osteoarthritis and neurodegenerative disorders.

Disorders related to organ transplantation

Some ORFX can be associated with disorders related to organ transplantation, in particular but not limited to organ rejection. Therapeutics of the invention, particularly those that modulate (or supply) activity, may be effective in treating or preventing diseases or disorders related to organ transplantation. Therapeutics of the invention (particularly Therapeutics that modulate the levels or activity of an aforementioned protein) can be assayed by any method known in the art for efficacy in treating or preventing such diseases and disorders related to organ transplantation. Such assays include *in vitro* assays for using cell culture models as described below, or *in vivo* assays using animal models of diseases and disorders related to organ transplantation, see *e.g.*, below. Potentially effective Therapeutics, for example but not by way of limitation, reduce immune rejection responses in animal models in comparison to controls.

Accordingly, once diseases and disorders related to organ transplantation are shown to be amenable to treatment by modulation of activity, such diseases or disorders can be treated or prevented by administration of a Therapeutic that modulates activity.

Cardiovascular Disease

GENX has been implicated in cardiovascular disorders, including in atherosclerotic plaque formation. Diseases such as cardiovascular disease, including cerebral thrombosis or hemorrhage, ischemic heart or renal disease, peripheral vascular disease, or thrombosis of other major vessel, and other diseases, including diabetes mellitus, hypertension, hypothyroidism, cholesterol ester storage disease, systemic lupus erythematosus, homocysteinemia, and familial protein or lipid processing diseases, and the like, are either directly or indirectly associated with atherosclerosis. Accordingly, Therapeutics of the invention, particularly those that modulate (or supply) activity or formation may be effective in treating or preventing

atherosclerosis-associated diseases or disorders. Therapeutics of the invention (particularly Therapeutics that modulate the levels or activity) can be assayed by any method known in the art, including those described below, for efficacy in treating or preventing such diseases and disorders.

5 A vast array of animal and cell culture models exist for processes involved in atherosclerosis. A limited and non-exclusive list of animal models includes knockout mice for premature atherosclerosis (Kurabayashi and Yazaki, 1996, *Int. Angiol.* 15: 187-194), transgenic mouse models of atherosclerosis (Kappel *et al.*, 1994, *FASEB J.* 8: 583-592), antisense oligonucleotide treatment of animal models (Callow, 1995, *Curr. Opin. Cardiol.* 10: 569-576),
10 transgenic rabbit models for atherosclerosis (Taylor, 1997, *Ann. N.Y. Acad. Sci.* 811: 146-152), hypercholesterolemic animal models (Rosenfeld, 1996, *Diabetes Res. Clin. Pract.* 30 Suppl.: 1-11), hyperlipidemic mice (Paigen *et al.*, 1994, *Curr. Opin. Lipidol.* 5: 258-264), and inhibition of lipoxygenase in animals (Sigal *et al.*, 1994, *Ann. N.Y. Acad. Sci.* 714: 211-224). In addition, *in vitro* cell models include but are not limited to monocytes exposed to low density lipoprotein
15 (Frostegard *et al.*, 1996, *Atherosclerosis* 121: 93-103), cloned vascular smooth muscle cells (Suttles *et al.*, 1995, *Exp. Cell Res.* 218: 331-338), endothelial cell-derived chemoattractant exposed T cells (Katz *et al.*, 1994, *J. Leukoc. Biol.* 55: 567-573), cultured human aortic endothelial cells (Farber *et al.*, 1992, *Am. J. Physiol.* 262: H1088-1085), and foam cell cultures (Libby *et al.*, 1996, *Curr Opin Lipidol* 7: 330-335). Potentially effective Therapeutics, for
20 example but not by way of limitation, reduce foam cell formation in cell culture models, or reduce atherosclerotic plaque formation in hypercholesterolemic mouse models of atherosclerosis in comparison to controls.

Accordingly, once an atherosclerosis-associated disease or disorder has been shown to be amenable to treatment by modulation of activity or formation, that disease or disorder can be
25 treated or prevented by administration of a Therapeutic that modulates activity.

Cytokine and Cell Proliferation/Differentiation Activity

A GENX protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered

to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D,
5 DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured by the following methods: Assays for T-cell or thymocyte proliferation include without limitation those described in: CURRENT PROTOCOLS IN IMMUNOLOGY, Ed by Coligan *et al.*, Greene Publishing
10 Associates and Wiley-Interscience (Chapter 3 and Chapter 7); Takai *et al.*, *J Immunol* 137:3494-3500, 1986; Bertagnoli *et al.*, *J Immunol* 145:1706-1712, 1990; Bertagnoli *et al.*, *Cell Immunol* 133:327-341, 1991; Bertagnoli, *et al.*, *J Immunol* 149:3778-3783, 1992; Bowman *et al.*, *J Immunol* 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or
15 thymocytes include, without limitation, those described by Kruisbeek and Shevach, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1, pp. 3.12.1-14, John Wiley and Sons, Toronto 1994; and by Schreiber, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan eds. Vol 1 pp. 6.8.1-8, John Wiley and Sons, Toronto 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells
20 include, without limitation, those described by Bottomly *et al.*, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto 1991; deVries *et al.*, *J Exp Med* 173:1205-1211, 1991; Moreau *et al.*, *Nature* 336:690-692, 1988; Greenberger *et al.*, *Proc Natl Acad Sci U.S.A.* 80:2931-2938, 1983; Nordan, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1 pp. 6.6.1-5, John Wiley and Sons,
25 Toronto 1991; Smith *et al.*, *Proc Natl Acad Sci U.S.A.* 83:1857-1861, 1986; Measurement of human Interleukin 11-Bennett, *et al.* In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto 1991; Ciarletta, *et al.*, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto 1991.

30 Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and

cytokine production) include, without limitation, those described in: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds., Greene Publishing Associates and Wiley-Interscience (Chapter 3, Chapter 6, Chapter 7); Weinberger *et al.*, *Proc Natl Acad Sci USA* 77:6091-6095, 1980; Weinberger *et al.*, *Eur J Immunol* 11:405-411, 1981; Takai *et al.*, *J Immunol* 137:3494-3500, 1986; Takai *et al.*, *J Immunol* 140:508-512, 1988.

Immune Stimulating or Suppressing Activity

A GENX protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), *e.g.*, in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (*e.g.*, HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases caused by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, *Leishmania* species, malaria species, and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, *i.e.*, in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to modulate immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response.

The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or energy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon re-exposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (*e.g.*, B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to energize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the

immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as described in Lenschow *et al.*, Science 257:789-792 (1992) and Turka *et al.*, Proc Natl Acad Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., FUNDAMENTAL IMMUNOLOGY, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and auto-antibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of auto-antibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythematosus in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., FUNDAMENTAL IMMUNOLOGY, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in vitro* activated T cells into the

patient. Another method of enhancing anti-vital immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be
5 capable of delivering a costimulatory signal to, and thereby activate, T cells *in vivo*.

In another application, up regulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (e.g., sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a
10 subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression
15 of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection *in vivo*.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells.
20 In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I α chain protein and β_2 microglobulin protein or an MHC class II α chain protein and an MHC class II β chain protein to thereby express MHC class I or MHC class II proteins on
25 the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a
30 B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor

specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods: Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Greene Publishing Associates and Wiley-Interscience (Chapter 3, Chapter 7); Herrmann *et al.*, *Proc Natl Acad Sci USA* 78:2488-2492, 1981; Herrmann *et al.*, *J Immunol* 128:1968-1974, 1982; Handa *et al.*, *J Immunol* 135:1564-1572, 1985; Takai *et al.*, *J Immunol* 137:3494-3500, 1986; Takai *et al.*, *J Immunol* 140:508-512, 1988; Herrmann *et al.*, *Proc Natl Acad Sci USA* 78:2488-2492, 1981; Herrmann *et al.*, *J Immunol* 128:1968-1974, 1982; Handa *et al.*, *J Immunol* 135:1564-1572, 1985; Takai *et al.*, *J Immunol* 137:3494-3500, 1986; Bowman *et al.*, *J Virology* 61:1992-1998; Takai *et al.*, *J Immunol* 140:508-512, 1988; Bertagnolli *et al.*, *Cell Immunol* 133:327-341, 1991; Brown *et al.*, *J Immunol* 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, *J Immunol* 144:3028-3033, 1990; and Mond and Brunswick In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, (eds.) Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Greene Publishing Associates and Wiley-Interscience (Chapter 3, Chapter 7); Takai *et al.*, *J Immunol* 137:3494-3500, 1986; Takai *et al.*, *J Immunol* 140:508-512, 1988; Bertagnolli *et al.*, *J Immunol* 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery *et al.*, *J Immunol* 134:536-544, 1995; Inaba *et al.*, *J Exp Med* 173:549-559, 1991; Macatonia *et al.*, *J Immunol* 154:5071-5079, 1995; Porgador *et al.*, *J Exp Med* 182:255-260, 1995; Nair *et al.*, *J Virol* 67:4062-4069, 1993; Huang *et al.*, *Science* 264:961-965, 1994; Macatonia *et al.*, *J Exp Med* 169:1255-1264, 1989; Bhardwaj *et al.*, *J Clin Investig* 94:797-807, 1994; and Inaba *et al.*, *J Exp Med* 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz *et al.*, *Cytometry* 13:795-808, 1992; Gorczyca *et al.*, *Leukemia* 7:659-670, 1993; Gorczyca *et al.*, *Cancer Res* 53:1945-1951, 1993; Itoh *et al.*, *Cell* 66:233-243, 1991; Zacharchuk, *J Immunol* 145:4037-4045, 1990; Zamai *et al.*, *Cytometry* 14:891-897, 1993; Gorczyca *et al.*, *Internat J Oncol* 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica *et al.*, *Blood* 84:111-117, 1994; Fine *et al.*, *Cell Immunol* 155: 111-122, 1994; Galy *et al.*, *Blood* 85:2770-2778, 1995; Toki *et al.*, *Proc Nat Acad Sci USA* 88:7548-7551, 1991.

Hematopoiesis Regulating Activity

A GENX protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, *e.g.* in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (*i.e.*, traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either in-vivo or ex-vivo (*i.e.*, in conjunction with bone marrow

transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

5 Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson *et al.* *Cellular Biology* 15:141-151, 1995; Keller *et al.*, *Mol. Cell.*
10 *Biol.* 13:473-486, 1993; McClanahan *et al.*, *Blood* 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, In: CULTURE OF HEMATOPOIETIC CELLS. Freshney, *et al.* (eds.) Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama *et al.*,
15 *Proc Natl Acad Sci USA* 89:5907-5911, 1992; McNiece and Briddeli, In: CULTURE OF HEMATOPOIETIC CELLS. Freshney, *et al.* (eds.) Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben *et al.*, *Exp Hematol* 22:353-359, 1994; Ploemacher, In: CULTURE OF HEMATOPOIETIC CELLS. Freshney, *et al.* eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Spoonceret *et al.*, In: CULTURE OF HEMATOPOIETIC CELLS. Freshney, *et al.*, (eds.) Vol pp.
20 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Sutherland, In: CULTURE OF HEMATOPOIETIC CELLS. Freshney, *et al.*, (eds.) Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

Tissue Growth Activity

A GENX protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for
25 wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.

A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation
30 employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation

induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendonitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein

may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, EPIDERMAL WOUND HEALING, pp. 71-112 (Maibach and Rovee, eds.), Year Book Medical

Publishers, Inc., Chicago, as modified by Eaglstein and Menz, *J. Invest. Dermatol* 71:382-84 (1978).

Activin/Inhibin Activity

A GENX protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin-b group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale *et al.*, *Endocrinology* 91:562-572, 1972; Ling *et al.*, *Nature* 321:779-782, 1986; Vale *et al.*, *Nature* 321:776-779, 1986; Mason *et al.*, *Nature* 318:659-663, 1985; Forage *et al.*, *Proc Natl Acad Sci USA* 83:3091-3095, 1986.

Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (*e.g.*, act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example,

attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population.

5 Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by following methods:

10 Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: CURRENT PROTOCOLS IN IMMUNOLOGY, Coligan *et al.*,
15 eds. (Chapter 6.12, MEASUREMENT OF ALPHA AND BETA CHEMOKINES 6.12.1-6.12.28); Taub *et al. J Clin Invest* 95:1370-1376, 1995; Lind *et al. APMIS* 103:140-146, 1995; Muller *et al., Eur J Immunol* 25: 1744-1748; Gruber *et al. J Immunol* 152:5860-5867, 1994; Johnston *et al., J Immunol* 153: 1762-1768, 1994.

Hemostatic and Thrombolytic Activity

20 A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for
25 treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (*e.g.*, stroke).

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assay for hemostatic and thrombolytic activity include, without limitation, those
30 described in: Linet *et al., J. Clin. Pharmacol.* 26:131-140, 1986; Burdick *et al., Thrombosis Res.*

45:413-419, 1987; Humphrey *et al.*, *Fibrinolysis* 5:71-79 (1991); Schaub, *Prostaglandins* 35:467-474, 1988.

Receptor/Ligand Activity

A protein of the present invention may also demonstrate activity as receptors, receptor
5 ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and
ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and
their ligands, receptor phosphatases and their ligands, receptors involved in cell—cell
interactions and their ligands (including without limitation, cellular adhesion molecules (such as
selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation,
10 antigen recognition and development of cellular and humoral immune responses). Receptors and
ligands are also useful for screening of potential peptide or small molecule inhibitors of the
relevant receptor/ligand interaction. A protein of the present invention (including, without
limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of
receptor/ligand interactions.

15 The activity of a protein of the invention may, among other means, be measured by the
following methods:

Suitable assays for receptor-ligand activity include without limitation those described in:
CURRENT PROTOCOLS IN IMMUNOLOGY, Ed by Coligan, *et al.*, Greene Publishing Associates and
Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions
20 7.28.1-7.28.22), Takai *et al.*, *Proc Natl Acad Sci USA* 84:6864-6868, 1987; Bierer *et al.*, *J. Exp.*
Med. 168:1145-1156, 1988; Rosenstein *et al.*, *J. Exp. Med.* 169:149-160 1989; Stoltenborg *et al.*,
J Immunol Methods 175:59-68, 1994; Stitt *et al.*, *Cell* 80:661-670, 1995.

Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The
25 anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the
inflammatory response, by inhibiting or promoting cell—cell interactions (such as, for example,
cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory
process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production
of other factors which more directly inhibit or promote an inflammatory response. Proteins
30 exhibiting such activities can be used to treat inflammatory conditions including chronic or acute

conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Tumor Inhibition Activity

In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth.

Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting

deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

Neural disorders in general include Parkinson's disease, Alzheimer's disease, Huntington's disease, multiple sclerosis, amyotrophic lateral sclerosis (ALS), peripheral neuropathy, tumors of the nervous system, exposure to neurotoxins, acute brain injury, peripheral nerve trauma or injury, and other neuropathies, epilepsy, and/or tremors.

10 EQUIVALENTS

From the foregoing detailed description of the specific embodiments of the invention, it should be apparent that particular novel compositions and methods involving nucleic acids, polypeptides, antibodies, detection and treatment have been described. Although these particular embodiments have been disclosed herein in detail, this has been done by way of example for purposes of illustration only, and is not intended to be limiting with respect to the scope of the appended claims that follow. In particular, it is contemplated by the inventors that various substitutions, alterations, and modifications may be made as a matter of routine for a person of ordinary skill in the art to the invention without departing from the spirit and scope of the invention as defined by the claims. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

Table 1

ORF#	Internal Identification Number	Protein similarity	Protein domain	Protein Classification	Cells or Tissues in which Gene is Expressed
1	13076366 (1, 2)	Novel Protein sim. GBank gij4691395[emb]CAB41562.1] - (AL049727) putative large secreted protein [Streptomyces coelicolor]		UNCLASSIFIED	264636
2	80248091 (3, 4)	Novel Protein sim. GBank gij2829508[sp]P71559[SUCC_MYCTU - SUCCINYL-COA SYNTHETASE BETA CHAIN (SCS-BETA)]	Contains protein domain (PF00549) - CoA-ligases	UNCLASSIFIED	264907, 264600, 264602, 264762, 264769, 264689, 264638, 264567
3	80415924 (5, 6)			UNCLASSIFIED	264910, 264604, 264634, 264805, 264636, 264691, 264907, 264692, 264629
4	82018837 (7, 8)			UNCLASSIFIED	264908, 264909, 264760, 264628, 264635
5	79970035 (9, 10)			UNCLASSIFIED	22279002, 264563
6	79842462 (11, 12)		Contains protein domain (PF00127) - Copper binding proteins, plastocyanin/azurin family	UNCLASSIFIED	264908
7	85515576 (13, 14)	Novel Protein sim. GBank gij4415926[gb]AAD20157] - (AC006282) unknown protein [Arabidopsis thaliana]		UNCLASSIFIED	20281099, 35696052, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264511, 285006, 264512, 285009, 264910, 264595, 264596, 264758, 264603, 264604, 264760, 264762, 264683, 264766, 264787, 264689, 35695917, 264690, 264692, 264693, 33657109, 264628, 264629, 35698423, 55811576, 35695855, 264630, 264631, 264632, 264634, 264638, 264637, 264638, 264639, 18108385, 264563, 264564, 264566, 264486
8	56924278 (15, 16)	Novel Protein sim. GBank gij58562[sp]O06459[NIRB_KLEPN - NITRITE REDUCTASE (NAD(P)H) LARGE SUBUNIT		reductase	264907
9	79394457 (17, 18)			UNCLASSIFIED	265007, 265019, 263872
10	7956459 (19, 20)			UNCLASSIFIED	264906
11	20414027 (21, 22)				264605
12	94141210 (23, 24)	Novel Protein sim. GBank gij3878145[emb]CAA9887.1] - (Z75543) similar to potassium channel protein [Caenorhabditis elegans]		misc_channel	264259, 265007, 83373044
13	20750551 (25, 26)			UNCLASSIFIED	264556, 264557, 264564
14	95105114 (27, 28)	Novel Protein sim. GBank gij2832781[emb]CAA12645] - (AJ225805) inward potassium channel alpha subunit [Egeria densa]	Contains protein domain (PF00023) - Ank repeat	potassium_channel	35696286, 35696052, 264510, 35695917, 264691, 264628, 35696423, 264555, 264558, 264559, 83373044
15	20458307 (29, 30)	Novel Protein sim. GBank gij1710781[sp]Q10234[RT05_SCHPO - PROBABLE MITOCHONDRIAL 40S RIBOSOMAL PROTEIN S5	Contains protein domain (PF00333) - Ribosomal protein S5	ribosomalprot	264604
16	20760358 (31, 32)				264555

17	20282744 (33, 34)	Novel Protein sim. GBank gll1174884[spIP44391]URE1 - UREASE ALPHA SUBUNIT (UREA AMIDOHYDROLASE)	Contains protein domain (PF00449) - Urease	264600	
18	80246804 (35, 36)	Novel Protein sim. GBank gll2281102 (AC002333) - SF16 Isolog [Arabidopsis thaliana]		26331827, 264555, 264557, 264638, 264558	
19	80076024 (37, 38)			22278996, 264907, 264910, 264600, 264693	
20	20724558 (39, 40)	Novel Protein sim. GBank gll2508112[spIP43872]UUP - ABC TRANSPORTER ATP-BINDING PROTEIN ULUP	UNCLASSIFIED transport	264602	
21	80417554 (41, 42)	Novel Protein sim. GBank gll1730203[spIP50442]GATM - RAT - GLYCINE AMIDINOTRANSFERASE PRECURSOR (L- ARGININE:GLYCINE AMIDINOTRANSFERASE) (TRANSAMIDINASE) (AT)	UNCLASSIFIED	22278995, 264906, 265008, 265010, 265011, 264602, 264605, 264768, 264688, 21808784, 264691, 18108376, 264638, 18108387, 264486	
22	11705858 (43, 44)			264885	
23	80419178 (45, 46)	Novel Protein sim. GBank gll1877329[embjCAB07071] - (Z92771) fadE25 [Mycobacterium tuberculosis]	Contains protein domain (PF00441) - Acyl-CoA dehydrogenase	264488, 264907, 264809, 264600, 264602, 264603, 264605, 264682, 264768, 32833988, 264638, 264488	
24	20291897 (47, 48)			264600	
25	80253774 (49, 50)			264593	
26	80255394 (51, 52)		UNCLASSIFIED	22278998, 58182435, 265018, 264586	
27	80235795 (53, 54)	Novel Protein sim. GBank gll408369[embjCAB42783.1] - (AL049841) putative 30S ribosomal protein S14 (Streptomyces coelicolor)	Contains protein domain (PF00253) - Ribosomal protein S14p/S29e	18108370, 35898423, 264635, 264555	
28	79483561 (55, 56)			264638	
29	82448765 (57, 58)	Novel Protein sim. GBank gll312280[spIO08333]K8PF - STRCO - 6- PHOSPHOFRUCTOKINASE (PHOSPHOFRUCTOKINASE) (PHOSPHOHEXOKINASE) (ATP-PFK)	UNCLASSIFIED kinase	264601, 264782, 264766, 264769, 264638	
30	79189333 (59, 60)		UNCLASSIFIED	264908, 265019, 264687, 21906764, 21908768	
31	19848158 (61, 62)		UNCLASSIFIED	264534	
32	82449495 (63, 64)	Novel Protein sim. GBank gll3560504 (AF027770) - unknown [Mycobacterium smegmatis]	UNCLASSIFIED	264905, 264605, 264762, 264766, 264687, 264689	
33	79582628 (65, 66)	Novel Protein sim. GBank gll2129003[pirj]G64507 - hypothetical protein MJ1865 - Methanococcus jannaschii	UNCLASSIFIED	264687	
34	87467657 (67, 68)		UNCLASSIFIED	60432289, 264600, 264602, 264760, 18108357, 264789, 265020, 264691, 264600, 264687, 264558, 264639	
35	95005170 (69, 70)	Novel Protein sim. GBank gll5420387[embjCAB46679.1] - (AJ243459) proteophosphoglycan [Leishmania major]	UNCLASSIFIED	264566	
36	19642042 (71, 72)	Novel Protein sim. GBank gll3287739[spIP73538]BIOB - BIOTIN SYNTHASE (BIOTIN SYNTHETASE)	synthase		
37	20369215 (73, 74)	Novel Protein sim. GBank gll231313[gbpAAD07128.1] - (AE000527) delta-1-pyrroline-5-carboxylate dehydrogenase [Helicobacter pylori 26695]	dehydrogenase	264603	

38	20466334 (75, 76)	Novel Protein sim. GBank gij3805970[embjCAA06231] - (A.004933) periplasmic nitrate reductase, large subunit [Rhodospseudomonas sp.]		reductase	264605
39	94300715 (77, 78)	Novel Protein sim. GBank gij1929448 (L63543) - endodermis [Xenopus laevis]	Contains protein domain (PF00207) - Alpha-2-macroglobulin family	complement	264905, 264806, 264807, 66712502, 264808, 264909, 264511, 265009, 264910, 55812038, 264758, 265011, 264762, 264882, 264763, 264764, 264766, 265022, 264893, 264828, 264831, 264634, 264635, 264555, 264638, 18108381, 264558, 18108385, 264482
40	20635625 (79, 80)			UNCLASSIFIED	264592
41	80023287 (81, 82)	Novel Protein sim. GBank gij854065[embjCAA58337] - (X63413) U88 [Human herpesvirus 6]			264591, 35695917
42	20724588 (83, 84)			UNCLASSIFIED	264602
43	20467069 (85, 86)	Novel Protein sim. GBank gij3820584 (AF086791) - carbamoylphosphate synthetase large subunit [Zymomonas mobilis]		synthase	264605
44	13085297 (87, 88)	Novel Protein sim. GBank gij2494784[spjQ50729]GUAA_MYCTU - GMP SYNTHASE (GLUTAMINE-HYDROLYZING) (GLUTAMINE AMIDOTRANSFERASE) (GMP SYNTHETASE)	Contains protein domain (PF00959) - GMP synthase C terminal domain	synthase	264769, 264638
45	39384711 (89, 90)	Novel Protein sim. GBank gij1881738 (U89888) - myosin-I binding protein Acan125 [Acanthamoeba castellanii]		UNCLASSIFIED	264769, 264510, 264508
46	95003398 (91, 92)			ngf	264568
47	11698624 (93, 94)			UNCLASSIFIED	264689
48	79407218 (95, 96)				18108385, 264635, 264828
49	21659844 (97, 98)			UNCLASSIFIED	264603
50	80503996 (99, 100)				264508, 264603, 264769, 264689, 264636, 264558, 264488
51	80265569 (101, 102)	Novel Protein sim. GBank gij3411177 (AF076240) - MocC [Rhizobium leguminosarum bv. viciae]		UNCLASSIFIED	264593, 18108387
52	79208528 (103, 104)	Novel Protein sim. GBank gij3914992[spjQ26284]SM41_HEMPU - 41 KD SPICULE MATRIX PROTEIN PRECURSOR (HSM41) (HPSMC)		struc	264634
53	36996970 (105, 106)	Novel Protein sim. GBank gij3880411 (AC004561) - putative proline-rich protein [Arabidopsis thaliana]		UNCLASSIFIED	264782
54	78570897 (107, 108)			UNCLASSIFIED	264630, 264909, 264766
55	80202703 (109, 110)	Novel Protein sim. GBank gij1633572 (U52064) - Herpesvirus salmori ORF73 homolog [Kaposi's sarcoma-associated herpes-like virus]			29331824, 264102, 265018, 18108376
56	8758408 (111, 112)	Novel Protein sim. GBank gij4321580[spjAAD15785] - (AF050114) alginate lyase [Pseudomonas sp. W7]			264604
57	11223386 (113, 114)		Contains protein domain (PF00076) - RNA recognition motif, (a.k.a. RRM, RBD, or RNP domain)	UNCLASSIFIED	264557

58	91227508 (115, 116)	Novel Protein sim. GBank gi 5616074 gb AAD45616.1 AF06194 - (AF061943) protease- derived STE20-like kinase PSK [Homo sapiens]	Contains protein domain (PF000069) - Eukaryotic protein kinase domain	kinase	58182573, 264259, 60432049, 35696032, 68712502, 264909, 265008, 265010, 265011, 264881, 29148764, 35695917, 60170615, 264891, 264692, 264693, 18108374, 35698423, 56182323, 60432113 264600, 264689, 264638
59	80077371 (117, 118)	Novel Protein sim. GBank gi 172920 sp P45630 RFE_MYCLE - PUTATIVE UNDECAPRENYL-PHOSPHATE ALPHA-N- ACETYLGLUCOSAMINYLTRANSFERASE	Contains protein domain (PF00953) - Glycosyl transferase	transferase	
60	12958341 (119, 120)	Novel Protein sim. GBank gi 1710216 (U79260) - unknown [Homo sapiens]			
61	80426808 (121, 122)			glycoprotein	264689 264766
62	13504968 (123, 124)				264630
63	16474553 (125, 126)			UNCLASSIFIED	265019
64	20724578 (127, 128)	Novel Protein sim. GBank gi 420845 pir JA47041 - transposase homolog (Insertion element (SAE1) - Alcaligenes eutrophus		UNCLASSIFIED	264602
65	79326308 (129, 130)	Novel Protein sim. GBank gi 3122312 sp O08134 KPYK_MYCTU - PYRUVATE KINASE (PK)	Contains protein domain (PF00224) - Pyruvate kinase	kinase	264563
66	46854384 (131, 132)	Novel Protein sim. GBank gi 3928723 emb CAA22218 - (AL034355) putative ABC transporter [Streptomyces coelicolor]		transport	22278996, 264558
67	78952543 (133, 134)	Novel Protein sim. GBank gi 231985 sp P30234 DHA_MYCTU - ALANINE DEHYDROGENASE (40 KD ANTIGEN)		dehydrogenase	265021
68	79817382 (135, 136)				264909
69	78841784 (137, 138)				264908
70	78871329 (139, 140)			UNCLASSIFIED	264908, 264908
71	85897456 (141, 142)				264908, 265021
72	87734977 (143, 144)	Novel Protein sim. GBank gi 4415926 gb AAD20157 - (AC006282) unknown protein [Arabidopsis thaliana]		UNCLASSIFIED	264488, 264905, 264906, 264907, 264908, 264511, 265008, 264910, 264758, 87188474, 264882, 264768, 264686, 264689, 35695917, 265021, 60170615, 264691, 35657023, 264692, 264693, 264629, 264631, 264639, 22278000
73	80025241 (145, 146)			UNCLASSIFIED	60424179, 264508, 264908, 265007, 264603, 264687, 264689, 264692, 18108387
74	20377410 (147, 148)			UNCLASSIFIED	264605
75	11619032 (149, 150)	Novel Protein sim. GBank gi 2853098 emb CAA16914 - (AL021767) vacuolar protein sorting [Schizosaccharomyces pombe]		UNCLASSIFIED	264689
76	95105303 (151, 152)	Novel Protein sim. GBank gi 4468811 emb CAB38212 - (AL035601) putative protein [Arabidopsis thaliana]		UNCLASSIFIED	83373044, 264906, 264557
77	10144718 (153, 154)	Novel Protein sim. GBank gi 854085 emb CAA58337 - (X83413) UBB [Human herpesvirus 6]		UNCLASSIFIED	264563
78	8758258 (155, 156)			UNCLASSIFIED	264604

79	94140190 (157, 158)	Novel Protein sim. GBank gl 5689453 db BAA83010.1 - (AB028981) KIAA1059 protein [Homo sapiens]	Contains protein domain (PF00169) - PH domain		35698266, 22278998, 29331822, 29331824, 29331825, 29331827, 264905, 264906, 264907, 86712502, 264908, 264909, 265008, 265009, 264910, 60170831, 55812036, 33109954, 285017, 265018, 284288, 284788, 56181582, 21906765, 21906768, 29148784, 265020, 264690, 264691, 264692, 264693, 60431528, 35896423, 264631, 264632, 264634, 264636, 264639, 83373044, 264564, 264566, 264567
80	82314840 (159, 160)		UNCLASSIFIED		264769, 264501, 265006, 264910, 264604, 264605, 264634, 264635, 264905, 264782, 264637, 264592, 264628, 264907, 264691, 264908, 264587, 264809, 264768
81	20467247 (161, 162)	Novel Protein sim. GBank gl 1723442 sp Q10258 YD2A_SCHPO - HYPOTHETICAL 69.0 KD PROTEIN C56F8.10 IN CHROMOSOME 1	reductase		264605
82	16331388 (163, 164)	Novel Protein sim. GBank gl 2895866 (AF045770) - methylmalonate semi-aldehyde dehydrogenase [Oryza sativa]	dehydrogenase		264587
83	94741180 (165, 166)	Novel Protein sim. GBank gl 3402673 (AC004697) - unknown protein [Arabidopsis thaliana]	UNCLASSIFIED		264488, 264508, 264509, 264905, 264908, 264909, 264511, 264591, 264593, 264594, 264595, 264598, 264758, 264603, 264780, 264681, 18108351, 264782, 264682, 264784, 264684, 264768, 264686, 264632, 264637, 264557, 264638, 264639, 18108385, 264566
84	80355375 (167, 168)	Novel Protein sim. GBank gl 173364 sp P45380 SAT1_RAT - SULFATE ANION TRANSPORTER 1 (CANALICULAR SULFATE TRANSPORTER) (SULFATE/CARBONATE ANTI-PORTER)	transport		264508, 264906, 264907, 264908, 264909, 264910, 264760, 264783, 264764, 264786, 264788, 264769, 35895855, 264636, 264637
85	80499600 (169, 170)	Novel Protein sim. GBank gl 2120988 pir J570682 - glycosyltransferase homolog - Bordelella pertussis	transferase		264605, 264762, 264687, 264769, 18108374, 264636, 264488
86	39559043 (171, 172)	Novel Protein sim. GBank gl 3256023 emb CAA17228.1 - (AL021897) hypothetical protein Rv1112 [Mycobacterium tuberculosis]			264910
87	13856808 (173, 174)		UNCLASSIFIED		264083

88	95344718 (175, 176)	Novel Protein sim. GBank gll559703(dhljBAA07552) - (D38549) ha1025 is new [Homo sapiens]		52644507, 52646385, 18108398, 65274572, 56182575, 56994075, 35696286, 22278997, 22278998, 22278999, 264092, 264093, 264094, 264095, 264259, 29331822, 29331824, 56182181, 66714117, 29331825, 29331826, 60432289, 29331827, 29331828, 35696052, 33658970, 264105, 264508, 264905, 264906, 264907, 264908, 26331830, 66712502, 52644045, 56182435, 265007, 265009, 264910, 60170831, 264592, 60431735, 60433356, 33657402, 264757, 60433438, 55812038, 264758, 21908754, 52646317, 33109954, 52644296, 67168474, 265011, 87188559, 264601, 265017, 265018, 264604, 265019, 264448, 264369, 264288, 264766, 52644228, 21906766, 21908787, 21906788, 21906789, 55811957, 35695917, 265020, 265021, 265022, 60170615, 52644150, 33657023, 65274620, 33657109, 27486281, 27486264, 33657349, 35695783, 264626, 263972, 18108374, 55810764, 35698423, 55811576, 65274791, 35695855, 60431850, 264638, 52644332, 56182323, 60170394, 83373044, 18108385, 18108387, 18108388, 56526486, 87168518, 60432113, 22278000, 22278002, 264482, 264584, 264488, 264600
89	80077389 (177, 178)	Novel Protein sim. GBank gll1710383[sp]P46352[RIPX, BACSU - PROBABLE INTEGRASE/RECOMBINASE RIPX		
90	82115999 (178, 180)	Novel Protein sim. GBank gll2498891[sp]P76403[YECCO, ECOLI - PUTATIVE PROTEASE IN BAER-GRK INTERGENIC REGION	UNCLASSIFIED	264760
91	78505950 (181, 182)	Novel Protein sim. GBank gll3367754[lemb]CAA20079] - (AL031155) hypothetical protein SC3A7.16c [Sireptomyces coelicolor]	protease	265008
92	79554871 (183, 184)	Novel Protein sim. GBank gll2895095 (AF011337) - putative E1-E2 ATPase [Mus musculus]	UNCLASSIFIED	264691
93	80496778 (185, 186)	Novel Protein sim. GBank gll1171919[sp]P46920[OPUA, BACSU - GLYCINE BETAINES TRANSPORT ATP-BINDING PROTEIN OPUA	ATPase-associated	264907, 264908, 264910, 265009, 264605, 264769
94	78646649 (187, 188)	Contains protein domain (PF00571) - transport CBS domain	transport	264906
95	11080238 (189, 190)			264594

98	94322125 (191, 192)	Novel Protein sim. GBank gl[4589560]dbj[BAA76802.1] - (AB023175) KIAA0958 protein [Homo sapiens]		UNCLASSIFIED	22278995, 22278999, 264259, 26331822, 29331826, 35698052, 29146499, 264509, 264908, 264907, 264908, 264909, 265007, 265008, 264910, 265009, 264593, 265010, 265017, 264604, 265019, 18108351, 264288, 264768, 264768, 264769, 21906765, 21906767, 21906769, 265020, 264692, 33657182, 35695763, 264628, 264629, 18108379, 264631, 264638, 18108381, 264559, 18108382, 83373044, 22278002, 264508
97	79605200 (193, 194)	Novel Protein sim. GBank gl[4583559]emb[CAB40388.1] - (AJ005255) OxyR [Erwinia chrysanthemi]		UNCLASSIFIED	
98	79427000 (195, 196)	Novel Protein sim. GBank gl[1001693]dbj[BAA10430] - (D64002) hypothetical protein [Synecocystis sp.]		UNCLASSIFIED	284909
99	20466524 (197, 198)	Novel Protein sim. GBank gl[1169479]sp[P43925]EFG_HAEN - ELONGATION FACTOR G (EF-G)		UNCLASSIFIED	284605
100	79840113 (199, 200)	Novel Protein sim. GBank gl[480897]pir[S37485 - gene msg1 protein - mouse		UNCLASSIFIED	264693
101	80203298 (201, 202)	Novel Protein sim. GBank gl[2694166]emb[CAA11773.1] - (AJ223998) PCZA361.18 [Amycolatopsis orientalis]		UNCLASSIFIED	265020, 264102, 263972
102	20467259 (203, 204)	Novel Protein sim. GBank gl[1731040]sp[P54509]YQH_H BACSU - HYPOTHETICAL HELICASE IN SINI-GCVT INTERGENIC REGION	Contains protein domain (PF00271) - Helicases conserved C-terminal domain	synthase	264605
103	20466368 (205, 206)	Novel Protein sim. GBank gl[854065]emb[CAA58337] - (X83413) U88 [Human herpesvirus 6]		UNCLASSIFIED	264591, 264595, 264602
104	80247572 (207, 208)	Novel Protein sim. GBank gl[1885117] (U70770) - furrowed		complement	264508
105	79605208 (209, 210)	Novel Protein sim. GBank gl[1885117] (U70770) - furrowed		UNCLASSIFIED	264511, 265009
106	28382058 (211, 212)	[Drosophila melanogaster]		ATPase associated	29331824, 284591, 21906754, 265019
107	80057781 (213, 214)	Novel Protein sim. GBank gl[1705505]sp[P54729]BS4 MOUSE - BS4 PROTEIN		transport	18108374, 35695917, 22278998, 264113, 264600, 264602, 264603, 265017, 264910, 264908, 264638, 264768
108	80237936 (215, 216)	Novel Protein sim. GBank gl[263577]emb[CAB15264] - (Z99120) similar to ABC transporter (ATP-binding protein) [Bacillus subtilis]	Contains protein domain (PF00005) - ABC transporter	UNCLASSIFIED	264758, 264603, 264630, 264638, 264637
109	95194148 (217, 218)	Novel Protein sim. GBank gl[233078]emb[CAB11265] - (Z98801) carboxypeptidase s precursor [Schizosaccharomyces pombe]			264687
110	79582823 (219, 220)				264584
111	39565458 (221, 222)			UNCLASSIFIED	264908
112	79560308 (223, 224)			UNCLASSIFIED	265007
113	17959439 (225, 226)			UNCLASSIFIED	264769
114	80502101 (227, 228)				

115	80251003 (228, 230)	Novel Protein sim. GBank gl 2246532 (U93872) - ORF 73, contains large complex repeat CR 73 (Kaposi's sarcoma-associated herpesvirus)		UNCLASSIFIED	52845156, 52845080, 33856970, 284592, 21908754, 27486284, 18108379, 35898423, 284635, 52844332, 18108382
116	81286889 (231, 232)				284905, 284906, 284907, 284908, 284909, 284910, 284758, 285010, 284763, 284802, 284784, 284766, 284885, 284886, 284788, 284789, 33857023, 284693, 33857109, 284628, 18108374, 284631, 284632, 284634, 284636, 284637, 284638, 284639, 58528488, 284585, 284586
117	78636695 (233, 234)			UNCLASSIFIED	284639, 284693
118	80222170 (235, 236)		Contains protein domain (PF00170) - bZIP transcription factor	UNCLASSIFIED	263974
119	91013071 (237, 238)	Novel Protein sim. GBank gl 732526 (U22327) - alpha2(IIV) collagen [Caenorhabditis elegans]		UNCLASSIFIED	22278996, 29331824, 60432289, 265007, 60433438, 284603, 284605, 18108351, 284769, 284689, 285020, 284534, 27486281, 284558, 83373044, 18108385, 284584
120	8756491 (239, 240)	Novel Protein sim. GBank gl 2131219 pir S50157 - cyclin-dependent kinase chain SRB10 - yeast (Saccharomyces cerevisiae)		kinase	284603
121	80028153 (241, 242)				284595
122	20457620 (243, 244)	Novel Protein sim. GBank gl 2052147 emb CAB08137 - (Z84752) ksgA [Mycobacterium tuberculosis]	Contains protein domain (PF00398) - Ribosomal RNA adenine dimethylases	transferase	284605
123	8758278 (245, 246)				284604
124	78104017 (247, 248)	Novel Protein sim. GBank gl 2833385 sp Q43134 UGST_SORBI - GRANULE-BOUND GLYCOPROTEIN (STARCH) SYNTHASE PRECURSOR		synthase	18108394, 18108397, 285008, 265007, 285008, 285010, 285011, 18108355, 18108379, 18108380, 18108384
125	87787988 (249, 250)	Novel Protein sim. GBank gl 475542 (U08255) - glutamate receptor delta-1 subunit [Rattus norvegicus]	Contains protein domain (PF00080) - Ligand-gated ion channel	misc_channel	284508, 284908, 285009, 284596, 22278002
126	58701283 (251, 252)	Novel Protein sim. GBank gl 5102785 emb CAB45200.1 - (AL079308) putative transcriptional regulator [Streptomyces coelicolor]			284511
127	20467287 (253, 254)			UNCLASSIFIED	284605
128	80248473 (255, 256)	Novel Protein sim. GBank gl 130120 sp P23820 PHOB_PSEAE - PHOSPHATE REGULON TRANSCRIPTIONAL REGULATORY PROTEIN PHOB	Contains protein domain (PF00072) - Response regulator receiver domain	phosphatase	284907, 284909, 284910, 284600, 284601, 284603, 284605, 18108351, 284693, 284557
129	85290543 (257, 258)	Novel Protein sim. GBank gl 2508493 sp P38036 YGCB_ECOLI - HYPOTHETICAL 100.5 KD PROTEIN IN IAP-CYSH INTERGENIC REGION	Contains protein domain (PF00270) - DEAD/DEAH box helicase	UNCLASSIFIED	35898423, 35898585, 284600, 284602, 284603, 284604, 284605, 284508, 284908, 284584, 284628, 284682, 284565, 284683
130	80085583 (259, 260)	Novel Protein sim. GBank gl 854065 emb CAA56337 - (X83413) U88 [Human herpesvirus 6]			284634
131	8495022 (261, 262)	Novel Protein sim. GBank gl 1076039 pir S54860 - ABC transporter PstC-2 chain - Mycobacterium tuberculosis	Contains protein domain (PF00526) - Binding-protein-dependent transport systems inner membrane component	transport	18108376, 284769, 28331826, 284689, 22278996, 285021, 284600, 284511, 284601, 284602, 284605, 284905, 284638

132	10887692 (263, 264)	Novel Protein sim. GBank gll1877340[emb CAB07068] - (Z92771) accA3 [Mycobacterium tuberculosis]	Contains protein domain (PF00289) - Carbamoyl-phosphate synthase (CPSase)	carboxylase	264836 264905, 264689
133	94630883 (265, 266)	Novel Protein sim. GBank gll1877340[emb CAB07068] - (Z92771) accA3 [Mycobacterium tuberculosis]			
134	79634880 (267, 268)	Novel Protein sim. GBank gll14585838[emb CAB40932.1] - (AL049630) putative NADH dehydrogenase [Streptomyces coelicolor]		dehydrogenase	264905, 264605, 265021
135	19885057 (269, 270)	Novel Protein sim. GBank gll1460074[emb CAB01049] - (Z77250) hypothetical protein Rv2568 [Mycobacterium tuberculosis]			264634
136	79846083 (271, 272)	Novel Protein sim. GBank gll2125896[emb CAA73511] - (Y13070) folypolyglutamate synthase [Streptomyces coelicolor]		synthase	264508
137	79619770 (273, 274)	Novel Protein sim. GBank gll5420387[emb CAB46679.1] - (AJ243459) proteophosphoglycan [Leishmania major]			264683, 264685, 264686, 264691, 264692, 264693
138	79635971 (275, 276)	Novel Protein sim. GBank gll5420387[emb CAB46679.1] - (AJ243459) proteophosphoglycan [Leishmania major]		UNCLASSIFIED	18108374, 18108385, 33657109, 33657182, 265010, 22278998, 265008, 265007, 285008, 285009, 284693
139	86688076 (277, 278)	Novel Protein sim. GBank gll5689912[emb CAB52075.1] - (AL109732) putative mutase [Streptomyces coelicolor. A3(?)]	Contains protein domain (PF01817) - Chorismate mutase	dehydrogenase	22278998, 265007, 264810, 60433356, 265010, 264602, 264605, 284768, 284688, 284769, 264693, 32833986, 18108374, 18108387
140	79825759 (279, 280)			UNCLASSIFIED	284908
141	20700094 (281, 282)				264600
142	80028104 (283, 284)	Novel Protein sim. GBank gll3581916[emb CAA20855] - (ALC31545) mutS family DNA mismatch repair protein [Schizosaccharomyces pombe]		nuclease	284602, 265017
143	11072274 (285, 286)			UNCLASSIFIED	264600
144	95009102 (287, 288)	Novel Protein sim. GBank gll334127[sp P97303 BAC2_MOUSE - TRANSCRIPTION REGULATOR PROTEIN BACH2 (BTB AND CNC HOMOLOG 2)]			263978, 264600, 264910, 284632, 264508, 264563, 264564, 284591, 284558, 264808, 284629, 284639
145	80027058 (289, 290)	Novel Protein sim. GBank gll3757569[emb CAA21315] - (AL031863) 1-evidence=predicted by content; 1-method=genefinder;084; 1-method_score=68.31; 1-evidence_end [Drosophila melanogaster]		UNCLASSIFIED	22278998, 284602
146	13085662 (291, 292)	Novel Protein sim. GBank gll140807[sp P24538 Y121_BURCE - INSERTION ELEMENT IS402 HYPOTHETICAL 24 KD PROTEIN]	Contains protein domain (PF01675) - Transposase		264687
147	94320366 (293, 294)	Novel Protein sim. GBank gll2827608[emb CAA16663] - (AL021648) uvrD2 [Mycobacterium tuberculosis]		helicase	264905, 264906, 264809, 264510, 265009, 60433358, 284600, 284601, 264604, 284605, 264687, 264769, 18108365, 85274791, 18108387
148	80248804 (295, 296)	Novel Protein sim. GBank gll2916947[emb CAA17385] - (AL021989) hypothetical protein Rv0986 [Mycobacterium tuberculosis]		transport	285009, 285010, 284600, 284602, 264603, 284604, 284605, 33657108, 284636

148	80249373 (297, 298)	Novel Protein sim. GBank gl11723073sp1Q110407081_MYCTU - HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN CY50.01	Contains protein domain (PF000005) - transport ABC transporter	265010, 264600, 264601, 264603, 264604, 27486285, 264836
150	20284748 (299, 300)	Novel Protein sim. GBank gl13724125emb1CAA118051 - (AJ224340) maltosephosphorylase [Lactobacillus sanfrancisco]		264600
151	20726398 (301, 302)	Novel Protein sim. GBank gl1728312sp1P07651DE08_ECOLI - PHOSPHOPENTOMUTASE (PHOSPHODEOXYRIBOMUTASE)	Contains protein domain (PF01676) - Metalloenzyme superfamily	264602
152	95002877 (303, 304)	Novel Protein sim. GBank gl12497952sp1P35667Y4TM_RHISN - HYPOTHETICAL HYDROLASE/PEPTIDASE Y4TM	peptidase	264602
153	80256665 (305, 306)	Novel Protein sim. GBank gl13123021sp1Q90508VIT1_FUNHE - VITELLOGENIN I PRECURSOR (VTG I) (CONTAINS: LIPOVITELLIN 1 (LV1); PHOSVITIN (PV); LIPOVITELLIN 2 (LV2))	UNCLASSIFIED	264593
154	82305968 (307, 308)	Novel Protein sim. GBank gl1418897ip1JN0443 - transcription initiation factor sigma homolog hrdB - Streptomyces aureofaciens	Contains protein domain (PF00140) - Sigma-70 factor	264910, 264762, 264691, 264634 264605
155	20429859 (309, 310)	Novel Protein sim. GBank gl1828710ip1J541739 - hypothetical protein - Escherichia coli	UNCLASSIFIED	264565
156	39564742 (311, 312)	Novel Protein sim. GBank gl13695013 (AF052586) - CirA [Pseudomonas aeruginosa]	Contains protein domain (PF00142) - 4Fe-4S iron sulfur cluster binding proteins, Nif-HfrC family	264691
157	10358887 (313, 314)	Novel Protein sim. GBank gl1073072ip1J55543 - cmlu protein - Pseudomonas syringae pv. syringae	UNCLASSIFIED	264805
158	78761938 (315, 316)	Novel Protein sim. GBank gl1173023sp1P48789RL30_STRCO - 50S RIBOSOMAL PROTEIN L30	Contains protein domain (PF00400) - WD domain, G-beta repeat	265008 264605
159	78890378 (317, 318)	Novel Protein sim. GBank gl1304869emb1CAB46028.11 - (AL031685) dJ963K23.2 (novel protein) [Homo sapiens]	Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger)	22278986, 264600, 264603, 35695917, 32833986, 35698423, 264638
160	11075119 (319, 320)	Novel Protein sim. GBank gl12661691emb1CAA157951 - (AL009204) putative protease [Streptomyces coelicolor]	interleukin	264112, 264532, 22279002
161	80055007 (321, 322)	Novel Protein sim. GBank gl14416478ip1AAD203781 - (AF125999) transposase [Mycobacterium avium]	UNCLASSIFIED	264639 264905, 264907, 264600
162	80016371 (323, 324)	Novel Protein sim. GBank gl176177ip1JQ0ECFT - hypothetical 38.8K protein (hs1 5' region) - Escherichia coli	UNCLASSIFIED	264605, 264486
163	11682306 (325, 326)		UNCLASSIFIED	265010
164	80077902 (327, 328)		UNCLASSIFIED	264600
165	10856057 (329, 330)		UNCLASSIFIED	264605, 264486
166	88095003 (331, 332)		UNCLASSIFIED	265010
167	16395460 (333, 334)		UNCLASSIFIED	264600
168	80078362 (335, 336)		UNCLASSIFIED	264556, 264557, 264558, 264559
169	80239581 (337, 338)			

170	78612364 (338, 340)	Novel Protein sim. GBank gl 140888 sp P27847 YIGK_ECOLI - HYPOTHETICAL 15.4 KD PROTEIN IN RECQ-PLD8 INTERGENIC REGION (F138)	Contains protein domain (PF01810) - LysE type translocator	284908 284595, 284604	
171	95293073 (341, 342)				
172	37787007 (343, 344)	Novel Protein sim. GBank gl 4210905 gb AAD12048.1 - (AF045609) AgIG [Sinorhizobium meliloti]	Contains protein domain (PF00528) - Binding-protein-dependent transport systems inner membrane component	284769	
173	57529860 (345, 346)	Novel Protein sim. GBank gl 132854 sp P02387 RL2_ECOLI - 50S RIBOSOMAL PROTEIN L2	Contains protein domain (PF00181) - Ribosomal Proteins L2	284769	
174	95293078 (347, 348)	Novel Protein sim. GBank gl 1881350 dbj BAA19377 - (AB001488) PROBABLE TRANSPORT PROTEIN, SIMILAR TO ANTIBIOTIC TRANSPORT-ASSOCIATED PROTEIN ACTII IN STREPTOMYCES COELICOLOR. [Bacillus subtilis]		284510, 284593, 284602, 284603, 284605, 284762, 284693	
175	79756270 (349, 350)	Novel Protein sim. GBank gl 2072722 emb CAB08328 - (Z95121) manA [Mycobacterium tuberculosis]	isomerase	284585	
176	80086898 (351, 352)	Novel Protein sim. GBank gl 1055198 (U40187) - similar to PIR-A41724 chicken LD (limb deformity) gene product and to formin; also P-rich region similar to collagen [Caenorhabditis elegans]	UNCLASSIFIED	284907, 284910, 284681, 284558	
177	86684852 (353, 354)	Novel Protein sim. GBank gl 2326738 emb CAB10952 - (Z98288) hypothetical protein Rv1695 [Mycobacterium tuberculosis]	Contains protein domain (PF01513) - Domain of unknown function	284768, 60424179, 284687, 284688, 284769, 29331828, 60432289, 18108376, 284889, 18108387, 32833986, 22278998, 265020, 284800, 284801, 284802, 284803, 284604, 284805, 284835, 284762, 284638, 284908, 284564, 284837, 284838, 284486, 60433358, 284768	
178	79559526 (355, 356)	Novel Protein sim. GBank gl 1906596 (U81788) - kinesin-73 [Drosophila melanogaster]	struct	284693, 33657109, 284635	
179	20263112 (357, 358)		UNCLASSIFIED	284593	
180	80488958 (359, 360)	Novel Protein sim. GBank gl 1169367 sp P45259 DNAB_HAEIN - REPLICATIVE DNA HELICASE	helicase	284769	
181	79585369 (361, 362)	Novel Protein sim. GBank gl 3170615 (AF059485) - DOC4 [Mus musculus]	UNCLASSIFIED	21806787, 284635, 284639, 18108384	
182	80577899 (363, 364)				
183	11614017 (365, 366)	Novel Protein sim. GBank gl 10766827 pr IS54172 - inorganic pyrophosphatase (EC 3.6.1.1) - common tobacco	UNCLASSIFIED	284259, 35886052, 56182435, 284511, 285018, 33657109, 284555, 284586 284690	
184	10174167 (367, 368)	Novel Protein sim. GBank gl 4371280 gb AAD18138 - (AC006260) hypothetical protein [Arabidopsis thaliana]	UNCLASSIFIED	284510	

185	21660822 (369, 370)	Novel Protein sim. GBank glj3008176[embjCAA18398.1] - (AL022304) putative mma transport regulator [Schizosaccharomyces pombe]		UNCLASSIFIED	264604
186	80070329 (371, 372)	Novel Protein sim. GBank glj2828802[spjP94408]YCLF_BACSU - HYPOTHETICAL 53.3 KD PROTEIN IN SFP-GERKA INTERGENIC REGION		transport	264595
187	80186611 (373, 374)	Novel Protein sim. GBank glj3150260[embjCAA19179] - (AL023834) cyclin [Schizosaccharomyces pombe]		UNCLASSIFIED	264369
188	20464942 (375, 376)	Novel Protein sim. GBank glj2145653[spjS72938] - hix protein - Mycobacterium leprae		Kinase	264605
189	82338215 (377, 378)	Novel Protein sim. GBank glj1881244[dbjBAA19271] - (AB001488) SIMILAR TO PYRUVATE OXIDASE AND ACETOLACTATE SYNTHASE. [Bacillus subtilis]	Contains protein domain (PF00205) - Thiamine pyrophosphate enzymes	UNCLASSIFIED	35686052, 264802, 264605, 264762, 264689, 35895917, 18108370, 18108372, 264636, 264585
190	80086821 (379, 380)	Novel Protein sim. GBank glj120226[spjP28725]FKBP_STRCH - FK506-BINDING PROTEIN (PEPTIDYL-PROLYL CIS-TRANS ISOMERASE) (PPIASE) (ROTAMASE)	Contains protein domain (PF00254) - FKBP-type peptidyl-prolyl cis-trans isomerases	synthase	264583
191	88095012 (381, 382)	Novel Protein sim. GBank glj4980892[gbjAAD35474.1]AE00171 - (AE001718) ABC transporter, ATP-binding protein [Thermotoga maritima]		isomerase	264508, 264604, 264805, 264769, 264555
192	16333378 (383, 384)	Novel Protein sim. GBank glj1705461[spjP53656]BIOA_ERWHE - ADENOSYLMETHIONINE-8-AMINO-7-OXONONANOATE AMINOTRANSFERASE (7,8-DIAMINO-PELARGONIC ACID AMINOTRANSFERASE) (DAPA AMINOTRANSFERASE)			264587
193	78810127 (385, 386)	Novel Protein sim. GBank glj1705461[spjP53656]BIOA_ERWHE - ADENOSYLMETHIONINE-8-AMINO-7-OXONONANOATE AMINOTRANSFERASE (7,8-DIAMINO-PELARGONIC ACID AMINOTRANSFERASE) (DAPA AMINOTRANSFERASE)		transport	264808, 264693
194	20464949 (387, 388)	Novel Protein sim. GBank glj1705461[spjP53656]BIOA_ERWHE - ADENOSYLMETHIONINE-8-AMINO-7-OXONONANOATE AMINOTRANSFERASE (7,8-DIAMINO-PELARGONIC ACID AMINOTRANSFERASE) (DAPA AMINOTRANSFERASE)			264805
195	13518369 (389, 390)	Novel Protein sim. GBank glj1705461[spjP53656]BIOA_ERWHE - ADENOSYLMETHIONINE-8-AMINO-7-OXONONANOATE AMINOTRANSFERASE (7,8-DIAMINO-PELARGONIC ACID AMINOTRANSFERASE) (DAPA AMINOTRANSFERASE)			264636
196	95005569 (391, 392)	Novel Protein sim. GBank glj1705461[spjP53656]BIOA_ERWHE - ADENOSYLMETHIONINE-8-AMINO-7-OXONONANOATE AMINOTRANSFERASE (7,8-DIAMINO-PELARGONIC ACID AMINOTRANSFERASE) (DAPA AMINOTRANSFERASE)	Contains protein domain (PF00202) - Aminotransferases class-III pyridoxal phosphate	gaba	264600, 264688, 264638
197	80248665 (393, 394)	Novel Protein sim. GBank glj312305[spjQ2778]K6PF_SCHMA - 6-PHOSPHOFRUCTOKINASE (PHOSPHOFRUCTOKINASE) (PHOSPHOHEXOKINASE)	Contains protein domain (PF00365) - Phosphofructokinase	kinase	264602, 264682, 264692, 18108374
198	78163635 (395, 396)	Novel Protein sim. GBank glj1781203[embjCAB08110] - (Z83859) gnd [Mycobacterium tuberculosis]			264638
199	78890715 (397, 398)	Novel Protein sim. GBank glj2842222 (AF030885) - telomere-associated recQ-like helicase [Ustilago maydis]	Contains protein domain (PF00393) - 6-phosphogluconate dehydrogenases		265008
200	78413848 (399, 400)	Novel Protein sim. GBank glj2894379[embjCAA7491.1] - (Y14573) ring finger protein [Hordeum vulgare]		UNCLASSIFIED	264595, 264596
201	86945924 (401, 402)	Novel Protein sim. GBank glj2894379[embjCAA7491.1] - (Y14573) ring finger protein [Hordeum vulgare]		UNCLASSIFIED	29331826, 265007, 264512, 33657402, 264598, 265017, 18108351, 264682, 264683, 264767, 264829, 55810764, 264634, 264635, 58182323, 60432113, 22278000

202	79588046 (403, 404)	Novel Protein sim. GBank gi 231772 sp P30598 CHS1 - USTMA - CHITIN SYNTHASE 1 (CHITIN-UDP ACETYL-GLUCOSAMINYL TRANSFERASE 1)	Contains protein domain (PF01644) - Chitin synthase	synthase	264600
203	79843927 (405, 406)	Novel Protein sim. GBank gi 1504042 dbj BAA13220 - (D86984) similar to yeast adenylate cyclase (S56776) (Homo sapiens)			22278995, 28331822, 29331825, 29331827, 264908, 21908754, 264683, 21908786, 21908769, 35998423, 264556
204	79855188 (407, 408)	Novel Protein sim. GBank gi 2633808 emb CAB13310 - (Z99111) similar to hypothetical proteins [Bacillus subtilis]		UNCLASSIFIED	264909
205	10090583 (409, 410)	Novel Protein sim. GBank gi 2134381 pir S60678 - polybromo 1 protein - chicken		transport	264909
206	8759473 (411, 412)	Novel Protein sim. GBank gi 2501040 sp O05814 SYP_MYCTU - PROLYL-TRNA SYNTHETASE (PROLINE-TRNA LIGASE) (PRORS)		UNCLASSIFIED	264604
207	20754522 (413, 414)	Novel Protein sim. GBank gi 2501040 sp O05814 SYP_MYCTU - PROLYL-TRNA SYNTHETASE (PROLINE-TRNA LIGASE) (PRORS)		UNCLASSIFIED	264556
208	20289261 (415, 416)	Novel Protein sim. GBank			264605
209	80071069 (417, 418)	Novel Protein sim. GBank			264605, 264689
210	80168800 (419, 420)	Novel Protein sim. GBank			264905, 264907, 264809, 264766, 264687, 264891, 264629, 18108374, 264638
211	80034539 (421, 422)	Novel Protein sim. GBank			263978
212	82442474 (423, 424)	Novel Protein sim. GBank gi 5031809 ref NP_005538.1 p ISLR - immunoglobulin superfamily containing leucine-rich repeat		UNCLASSIFIED	264508, 264905, 264906, 264907, 264908, 264600, 264762, 264534, 264632, 264634, 264835, 264639, 264486
213	80248562 (425, 426)	Novel Protein sim. GBank gi 3122359 sp O33123 LEU2 MYCLE - 3 - ISOPROPYL MALATE DEHYDRATASE LARGE SUBUNIT (ISOPROPYL MALATE ISOMERASE) (ALPHA-IPM ISOMERASE) (IPMI)	Contains protein domain (PF00330) - Aconitase family (aconitase hydratase)	isomerase	22278998, 284508, 264600, 264602, 264603, 284505, 33557023, 264565, 264486
214	80079381 (427, 428)	Novel Protein sim. GBank gi 116238 sp P19421 CH60 COXBU - 60 KD CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN) (HEAT SHOCK PROTEIN B)	Contains protein domain (PF00118) - TCP-1/cpn60 chaperonin family	eph	264600, 264693
215	14973283 (429, 430)	Novel Protein sim. GBank			
216	80177716 (431, 432)	Novel Protein sim. GBank gi 2506924 sp P49754 VP41_HUMAN - VACUOLAR ASSEMBLY PROTEIN VPS41 HOMOLOG (S53)	Contains protein domain (PF00098) - Zinc finger, C2H2 type	UNCLASSIFIED	264629
217	79603634 (433, 434)	Novel Protein sim. GBank		dna_rna_bind	264448
218	80258475 (435, 436)	Novel Protein sim. GBank gi 1173288 sp P38106 IRSEA, ECOLI - SIGMA-E FACTOR NEGATIVE REGULATORY PROTEIN			264508
219	20438797 (437, 438)	Novel Protein sim. GBank gi 1781097 emb CAB06231 - (Z83864) glbB [Mycobacterium tuberculosis]		mapolymerase	264594
220	13489572 (439, 440)	Novel Protein sim. GBank gi 2984703 (AF052427) - unknown [Trypanosoma cruzi]		synthase	264604
221	11287498 (441, 442)	Novel Protein sim. GBank gi 4567313 dbj BAA76708.1 - (AB025248) alpha-1,2-mannosidase [Bacillus sp. M-90]		nucleaseinhib	264689
				UNCLASSIFIED	264555

222	79862802 (443, 444)	Novel Protein sim. GBank gi 1877268 emb CAB07049 - (Z92770) hypothetical protein Rv0143c [Mycobacterium tuberculosis]		UNCLASSIFIED	284605, 284769, 35896423
223	83053869 (445, 446)			UNCLASSIFIED	284808, 284907, 284603
224	79557920 (447, 448)			UNCLASSIFIED	284894, 284893
225	79558541 (449, 450)	Novel Protein sim. GBank gi 2274851 db BAA21515 - (D64159) 3-7 gene product [Homo sapiens]		UNCLASSIFIED	284692
226	79172397 (451, 452)	Novel Protein sim. GBank gi 868245 (U9488) - C58C10.7 gene product [Caenorhabditis elegans]		UNCLASSIFIED	22278998, 284112, 33657023, 263881
227	81777198 (453, 454)			UNCLASSIFIED	35695917, 284636, 284907
228	79872285 (455, 456)				284768, 284907, 284908, 284892, 284593, 284639
229	79838268 (457, 458)				284908, 284910
230	11013209 (459, 460)			UNCLASSIFIED	284631
231	20622207 (461, 462)	Novel Protein sim. GBank gi 1835114 emb CAA71733 - (Y10744) homoserine O-acetyltransferase [Leptospira meyeri]			284908, 284600, 284603, 284692
232	80055035 (463, 464)			UNCLASSIFIED	284600, 284603, 284605, 284687, 284789
233	80063054 (465, 466)	Novel Protein sim. GBank gi 2842340 (AF032970) - imidazolone propionate hydrolase [Pseudomonas putida]	Contains protein domain (PF00449) - Urease		284604
234	7523998 (467, 468)	Novel Protein sim. GBank gi 3510505 (AF030881) - polypeptide [Fugu rubripes]		UNCLASSIFIED	284369
235	80203671 (469, 470)			UNCLASSIFIED	284108
236	78940001 (471, 472)	Novel Protein sim. GBank gi 2104609 emb CAB08805 - (Z95398) PckA [Mycobacterium leprae]		carboxylase	284805
237	11755273 (473, 474)				284681
238	78461401 (475, 476)			UNCLASSIFIED	284639
239	82435180 (477, 478)	Novel Protein sim. GBank gi 2495617 sp Q57252 YDIJ_HAEIN - HYPOTHETICAL PROTEIN H11163	Contains protein domain (PF00037) - 4Fe-4S ferredoxins and related iron-sulfur cluster binding domains.		284808, 285010, 284603, 284762, 284682, 284636, 284638, 284486
240	21635575 (479, 480)	Novel Protein sim. GBank gi 3183458 sp P75786 YLIA_ECOLI - HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN YLIA	transport		284259, 284769
241	80377307 (481, 482)	Novel Protein sim. GBank gi 3875920 emb CAB04111 - (Z81503) predicted using GeneFinder: similar to collagen; cDNA EST EMBL:D65450 comes from this gene; cDNA EST EMBL:D68888 comes from this gene [Caenorhabditis elegans]		UNCLASSIFIED	284808, 284809, 284764, 284639
242	82148454 (483, 484)			UNCLASSIFIED	284489, 284907, 284808, 284511, 284760, 284784, 284892, 284635, 284637
243	79633207 (485, 486)			UNCLASSIFIED	284908
244	80248682 (487, 488)	Novel Protein sim. GBank gi 2624302 emb CAA15575 - (AL008987) aid [Mycobacterium tuberculosis]	dehydrogenase		284600, 284602, 284605, 284769, 284689
245	79863543 (489, 490)	Novel Protein sim. GBank gi 2920625 (AF044499) - vgrE protein [Escherichia coli]		UNCLASSIFIED	284907, 284758
246	79162928 (491, 492)	Novel Protein sim. GBank gi 5420387 emb CAB46679.1 - (AJ243459) proteophosphoglycan [Leishmania major]	Contains protein domain (PF01106) - NIDU-like domain		284637, 18108381, 18108387, 284555

247	79873185 (493, 494)	Novel Protein sim. GBank gij183908[embjCAB06848] - (Z85982) argB [Mycobacterium tuberculosis]		kinase	284809, 284691, 35696423, 18108387
248	80488983 (495, 496)	Novel Protein sim. GBank gij1168574[spP42464]ATPB_CORGL - ATP SYNTHASE BETA CHAIN		synthase	35698286, 284907, 284511, 284602, 284768, 284688, 285021, 35695855, 18108385
249	79764645 (497, 498)			UNCLASSIFIED	284907, 284910, 285011, 284762, 284638
250	79619980 (499, 500)				21908766, 284692
251	84359488 (501, 502)			UNCLASSIFIED	52645158, 29331822, 29331824, 52644045, 285018, 21908765, 21908768, 285020, 27488261, 27488265, 35695763, 18108376, 284556, 284559, 284565
252	79737758 (503, 504)	Novel Protein sim. GBank gij327166[dbjBAA31651] - (AB014578) KIAA0678 protein [Homo sapiens]			284685, 284687, 284632
253	20443124 (505, 506)	Novel Protein sim. GBank gij303688[embjCAA18513] - (AL022374) putative ATP-dependent DNA helicase [Streptomyces coelicolor]		helicase	284604
254	80027421 (507, 508)	Novel Protein sim. GBank gij3915488[spO34961]YJMB_BACSU - HYPOTHETICAL SYMPORTER IN COTI-RAPA INTERGENIC REGION		UNCLASSIFIED	284508, 284806, 284602, 284687, 285021, 284488
255	11398315 (509, 510)	Novel Protein sim. GBank gij1685720[dbjBAA04134] - (D17312) diarrheal toxin [Bacillus cereus]		UNCLASSIFIED	284593
256	80028158 (511, 512)	Novel Protein sim. GBank gij465787[spP34422]YL31_CAEEL - HYPOTHETICAL 86.0 KO PROTEIN F4489.1 IN CHROMOSOME III	Contains protein domain (PF00328) - Prolyl oligopeptidase family	peptidase	284602, 284692
257	20289282 (513, 514)	Novel Protein sim. GBank gij1172039[spP42315]SCOA_BACSU - PROBABLE SUCCINYL-COA:3-KETOACID-COENZYME A TRANSFERASE SUBUNIT A (SUCCINYL COA:3-OXOACID COA-TRANSFERASE) [OXCT A]	Contains protein domain (PF01144) - Coenzyme A transferase	transferase	284605
258	20459484 (515, 516)	Novel Protein sim. GBank gij3127836[embjCAA18902] - (AL023496) hypothetical protein [Streptomyces coelicolor]		UNCLASSIFIED	284604
259	79910152 (517, 518)			collagen	284681, 284686, 284692
260	20378437 (519, 520)			UNCLASSIFIED	284692, 284556
261	20285883 (521, 522)	Novel Protein sim. GBank gij123781[spP24221]HUTH_STRGR - HISTIDINE AMMONIA-LYASE (HISTIDASE)	Contains protein domain (PF00221) - Phenylalanine and histidine ammonia lyases	UNCLASSIFIED	284600
262	80189317 (523, 524)			UNCLASSIFIED	285017, 284369
263	80095045 (525, 526)	Novel Protein sim. GBank gij3924708[embjCAA84648] - (Z35597) Weak similarity with sea squirt nitrogen precursor protein (blastp score 71); cDNA EST EMBL: T02069 comes from this gene; cDNA EST EMBL: D76135 comes from this gene; cDNA EST EMBL: D73147 comes from this gene; cDNA EST EMBL: ...		UNCLASSIFIED	284488, 284905, 284906, 284907, 284908, 284909, 284512, 284910, 284756, 284598, 284604, 285019, 284805, 284760, 18108351, 284763, 284764, 284288, 284768, 284769, 284789, 284691, 284692, 284693, 284628, 284634, 284635, 284555, 284636, 284638, 284639
264	87370828 (527, 528)	Novel Protein sim. GBank gij3043734[dbjBAA25531] - (AB011177) KIAA0805 protein [Homo sapiens]	Contains protein domain (PF00047) - Immunoglobulin domain	protease	284259, 284908, 21808754, 285018, 285019, 285020

265	85355646 (529, 530)	Novel Protein sim. GBank gi 4589624 dbj BAA78834.1 - (AB023207) KIAA0980 protein [Homo sapiens]	kinase	264488, 35898286, 28331824, 56182181, 35898052, 284508, 284905, 284908, 284907, 68712502, 284908, 284909, 284511, 284512, 284910, 284592, 284995, 284758, 284598, 55811386, 284600, 285017, 284603, 284604, 284605, 284780, 18108351, 284782, 284881, 284784, 284288, 284788, 284788, 284789, 21908765, 21908767, 21908769, 285020, 284691, 33857023, 33857109, 33857182, 284828, 35898423, 35898555, 284630, 284631, 284632, 284634, 284635, 284638, 284555, 284638, 83373044, 56528488, 87188518, 284584, 284588, 284488
266	79588075 (531, 532)		UNCLASSIFIED	284600
267	11382222 (533, 534)		UNCLASSIFIED	284828
268	79809598 (535, 536)		UNCLASSIFIED	284687, 284769, 284689
269	80025810 (537, 538)		UNCLASSIFIED	284602
270	84381144 (539, 540)	Novel Protein sim. GBank gi 4507387 ref NP_003182.1 pTARS - threonyl-tRNA synthetase	UNCLASSIFIED	284693
271	79552301 (541, 542)	Novel Protein sim. GBank	UNCLASSIFIED	284908, 284693
272	9874778 (543, 544)	gi 4980738 gb AAD5331.1 AE001707 glucose- 1-phosphate adenylyltransferase [Thermotoga maritima]	synthase	284908
273	12840694 (545, 546)	Novel Protein sim. GBank gi 1168224 sp P44589 SNTD_HAEIN - PROBABLE 5'- NUCLEOTIDASE PRECURSOR	UNCLASSIFIED	284688
274	38524248 (547, 548)			284584
275	82787041 (549, 550)	Novel Protein sim. GBank gi 3253159 (AF005355) - translation initiation factor eIF2C [Oryctolagus cuniculus]	UNCLASSIFIED	284907, 284908, 284909, 284786, 284788, 284691, 284632, 284638
276	86671073 (551, 552)	Novel Protein sim. GBank gi 134920 sp P21997 SSGP_VOLCA - SULFATED SURFACE GLYCOPROTEIN 185 (SSG 185)		285008, 60432229
277	80078735 (553, 554)	Novel Protein sim. GBank gi 128021 sp P20984 OBG_BACSU - SPO0B-ASSOCIATED GTP-BINDING PROTEIN	ribosomal/prot	284600, 18108387
278	12866947 (555, 556)		UNCLASSIFIED	284689
279	95292719 (557, 558)	Novel Protein sim. GBank gi 78839 pir S03812 - uvrB protein - Micrococcus luteus	nuclease	284508, 284604, 21906764, 284638, 284557, 284404
280	5603617 (559, 560)			284259
281	80249599 (561, 562)	Novel Protein sim. GBank gi 3123160 sp Q18984 YLN2 CAEEL - HYPOTHETICAL 48.2 KD TRP-ASP REPEATS CONTAINING PROTEIN D2013.2 IN CHROMOSOME II		18108392, 284634, 284555, 284556, 284557, 284558
282	18598882 (563, 564)		UNCLASSIFIED	285019
283	20814211 (565, 566)		UNCLASSIFIED	284553

284	91212160 (587, 588)	Novel Protein sim. GBank gl 2429094 (U58632) - acetyl xylan esterase; AxeA [Thermotoga neopolitana]	Contains protein domain (PF00300) - Phosphoglycerate mutase family	UNCLASSIFIED	35696052, 29331828, 264508, 284905, 264600, 264602, 284605, 284682, 284764, 58181562, 21906764, 18108378, 264638, 284559, 18108387
285	8757940 (569, 570)	Novel Protein sim. GBank gl 2072674 (emb CAB08305) - (Z95120) rhlE [Mycobacterium tuberculosis]	Contains protein domain (PF00270) - DEAD/DEAH box helicase	UNCLASSIFIED	264603
286	80503235 (571, 572)	Novel Protein sim. GBank gl 765323 (bbs157676 - (S74439) silk fibroin heavy chain (C-terminal) [Bombyx mori=silkworms, Peptide Partial, 633 aa] [Bombyx mori])		ATPase-associated	35696052, 264769, 264638
287	12745521 (573, 574)	Novel Protein sim. GBank gl 765323 (bbs157676 - (S74439) silk fibroin heavy chain (C-terminal) [Bombyx mori=silkworms, Peptide Partial, 633 aa] [Bombyx mori])		UNCLASSIFIED	264689
288	20756502 (575, 578)	Novel Protein sim. GBank gl 1870009 (emb CAB08860) - (Z92539) hypothetical protein Rv1019 [Mycobacterium tuberculosis]	Contains protein domain (PF00440) - Bacterial regulatory proteins, telR family	collagen	284557
289	80043804 (577, 578)	Novel Protein sim. GBank gl 1870009 (emb CAB08860) - (Z92539) hypothetical protein Rv1019 [Mycobacterium tuberculosis]			264593, 264600
290	80430175 (579, 580)	Novel Protein sim. GBank gl 2506684 (sp P40120) YDCQ_ECOLI - 59.4 PROTEIN IN TRG-RIML INTERGENIC REGION PRECURSOR		UNCLASSIFIED	264768
291	20747431 (581, 582)	Novel Protein sim. GBank gl 625182 (L39015) - mitochondrial glutamyl-tRNA synthetase [Saccharomyces cerevisiae]		UNCLASSIFIED	284601
292	80052555 (583, 584)	Novel Protein sim. GBank gl 1718065 (sp P5328) UVRD_MYCLE - PUTATIVE DNA HELICASE II HOMOLOG		UNCLASSIFIED	264605
293	80062519 (585, 586)	Novel Protein sim. GBank gl 1718065 (sp P5328) UVRD_MYCLE - PUTATIVE DNA HELICASE II HOMOLOG	helicase		264909, 264605, 264687, 264689, 284692
294	79830303 (587, 588)	Novel Protein sim. GBank gl 117422 (sp P10040) CRB_DROME - CRUMBS PROTEIN PRECURSOR (95F)	Contains protein domain (PF00008) - EGF-like domain	oncogene	35696052, 264906, 265011, 284628, 55811576
295	79444180 (589, 590)	Novel Protein sim. GBank gl 1181619 (dbj BAA11565) - (D82384) a variant of TSC-22 [Gallus gallus]			52644507, 29331822, 264592, 265020, 264639
296	79607076 (591, 592)	Novel Protein sim. GBank gl 3649789 (dbj BAA33403) - (AB012228) SecA [Vibrio alginolyticus]		synthase	264508
297	79631287 (593, 594)	Novel Protein sim. GBank gl 5689967 (emb CAB52004.1) - (AL109663) putative membrane protein [Streptomyces coelicolor A3(2)]		UNCLASSIFIED	264905, 264687, 264638
298	80418898 (595, 596)			UNCLASSIFIED	264905, 264691, 264639, 264766

299	95283288 (597, 598)	Novel Protein sim. GBank gl 220637 dbj BAA01477 - (D10627) zinc finger protein [Mus musculus]	Contains protein domain (PF00088) - Zinc finger, C2H2 type	264488, 263994, 56994075, 22278997, 22278998, 22278999, 20281099, 29331824, 29331825, 29331826, 60432289, 29331827, 29331828, 264905, 264906, 264907, 264908, 52644045, 264909, 264511, 265008, 264910, 264595, 264598, 264758, 33657084, 87168559, 265018, 265019, 264784, 264288, 264788, 264687, 58181582, 264789, 21906785, 21906788, 21906789, 33657023, 264692, 33657109, 27485281, 18108370, 264628, 264629, 55811578, 35695855, 264631, 264634, 264635, 264638, 264639, 83373044, 18108387, 87168518, 22278000, 22279002, 264585, 264586, 264587
300	20711340 (599, 600)	Novel Protein sim. GBank gl 145922 (M20981) - Iron diclitate transport protein precursor [Escherichia coli]	UNCLASSIFIED	264602
301	13511332 (601, 602)	Novel Protein sim. GBank	transport	264687
302	9875260 (603, 604)	Novel Protein sim. GBank gl 1174881 sp P44594 TGT_HAEIN - QUEUINE TRNA- RIBOSYL TRANSFERASE (TRNA-GUANINE TRANSGLYCOSYLASE) (GUANINE INSERTION ENZYME)		264808
303	79574895 (605, 606)	Novel Protein sim. GBank gl 67985 pir JHUNVAV - helicase (EC 3.6.1.-) - Autographa californica nuclear polyhedrosis virus		264889
304	20711344 (607, 608)	Novel Protein sim. GBank	helicase	264602
305	80412520 (609, 610)	Novel Protein sim. GBank gl 728867 sp P40602 APG_ARATH - ANTER-SPECIFIC PROLINE-RICH PROTEIN APG PRECURSOR		264783
306	8515876 (611, 612)	Novel Protein sim. GBank gl 1657554 gb AAB18082.1 - (U73857) hypothetical protein [Escherichia coli]	UNCLASSIFIED	283978
307	80222801 (613, 614)	Novel Protein sim. GBank	UNCLASSIFIED	265010, 21906768, 265020, 18108374, 283977
308	80084305 (615, 616)	Novel Protein sim. GBank gl 1710612 sp Q10783 RNH2_MYCTU - PROBABLE RIBONUCLEASE HII (RNASE HII)	Contains protein domain (PF01351) - Ribonuclease HII	284910, 264600, 264605, 264687, 264689, 264638, 18108387
309	80504138 (617, 618)	Novel Protein sim. GBank gl 5420387 emb CAB46679.1 - (AJ243459) proteophosphoglycan [Leishmania major]		264769
310	80053618 (619, 620)	Novel Protein sim. GBank gl 1144522 (U34957) - phosphoribosylaminimidazole succinocarboxamide synthase [Mycobacterium tuberculosis]	synthase	264603
311	11090659 (621, 622)	Novel Protein sim. GBank gl 1144522 (U34957) - phosphoribosylaminimidazole succinocarboxamide synthase [Mycobacterium tuberculosis]		284602
312	80054347 (623, 624)	Novel Protein sim. GBank	UNCLASSIFIED	264586
313	80046188 (625, 626)	Novel Protein sim. GBank		284603, 264587

314	87845112 (627, 628)	Novel Protein sim. GBank glj3661583 (AF092175) - Ikaros [Dario reflo]	Contains protein domain (PF00320) - GATA zinc finger	dna_rna_bind	264259, 60432289, 28331828, 264903, 264908, 264908, 264909, 265008, 264910, 60432229, 33657402, 60433438, 33109954, 265011, 265017, 264603, 265018, 264288, 264768, 264892, 35895763, 264628, 264629, 264639, 60170394, 22279002, 264568, 264508, 264600, 264762, 264687, 264768, 52644229, 264769, 264689, 264635, 264636, 264638, 264488
315	82356091 (629, 630)	Novel Protein sim. GBank glj1852820 [dbj] [BAA17540] - (D80907) pyridine nucleotide transhydrogenase beta subunit [Synechocystis sp.]		UNCLASSIFIED	
316	78911071 (631, 632)	Novel Protein sim. GBank glj118244 [sp] [P24176] [DAPE, ECOLI - SUCCINYL-DIAMINOPIMELATE DESUCCINYLASE (SDAP)]		UNCLASSIFIED	264605
317	20466944 (633, 634)	Novel Protein sim. GBank glj4680229 [gb] [AAD27583.1] [AF118274] DNB-5 [Homo sapiens]	Contains protein domain (PF00528) - Dictyostellum (slime mold) repeats	transport	264908, 264909, 264810, 264593, 264594, 264760, 264288, 264768, 264769, 21908769, 264691, 264693, 264628, 65274791, 264635, 264636, 264638, 83373044, 22279002, 264568
318	94141836 (635, 636)	Novel Protein sim. GBank glj1149693 [emb] [CAA60220] - (X86499) rbsC [Clostridium perfringens]		transport	265018
319	17289360 (637, 638)	Novel Protein sim. GBank glj11033 [sp] [O05314] [GLGC, MYCTU - GLUCOSE-1-PHOSPHATE ADENYLTRANSFERASE (ADP-GLUCOSE SYNTHASE) (ADP-GLUCOSE PYROPHOSPHORYLASE)]		synthase	264687
320	13527675 (639, 640)	Novel Protein sim. GBank glj1600716 (U68234) - all-trans-retinoic acid 4-hydroxylase [Dario reflo]		cyto450	264509, 264908, 264907, 264908, 265009, 264596, 264764, 264628, 264634, 264635, 264638, 264639, 83373044, 264587
321	94134387 (641, 642)	Novel Protein sim. GBank glj1160355 (U33058) - UNC-89 [Caenorhabditis elegans]		UNCLASSIFIED	55811150, 264681, 60431528, 55810764
322	66489053 (643, 644)			UNCLASSIFIED	264488, 265009, 264593, 264628, 264635
323	94653725 (645, 646)			UNCLASSIFIED	264687
324	78174383 (647, 648)			UNCLASSIFIED	264693
325	78862691 (649, 650)			UNCLASSIFIED	264288, 18108385
326	28774974 (651, 652)	Novel Protein sim. GBank glj451544 (U04267) - proline-rich cell wall protein [Gossypium barbadense]			264488, 264905, 264509, 264910
327	79776267 (653, 654)			UNCLASSIFIED	264592
328	80253202 (655, 656)			UNCLASSIFIED	264510
329	10173821 (657, 658)			UNCLASSIFIED	264259, 264908
330	86597767 (659, 660)	Novel Protein sim. GBank glj4191358 (AF087825) - claudin-7 [Mus musculus]			
331	79754888 (661, 662)	Novel Protein sim. GBank glj80741 [pir] [S20912 - regulatory protein wh1B - Streptomyces coelicolor]		transcriptfactor	264910, 264687, 264689, 264636, 264567
332	80071440 (663, 664)	Novel Protein sim. GBank glj114049 [sp] [P19480] [AHPF, SALTY - ALKYL HYDROPEROXIDE REDUCTASE SUBUNIT F (ALKYL HYDROPEROXIDE REDUCTASE F52A PROTEIN)]		reductase	35698423, 264636, 264638, 264565
333	13009555 (665, 666)				264687

334	80230771 (667, 668)	Novel Protein sim. GBank gl 322228 pir S32227 - glutamate dehydrogenase (NADP+) (EC 1.4.1.4) - Corynebacterium glutamicum	Contains protein domain (PF00208) - Glutamate/Leucine/Phenylalanine/Va line dehydrogenase	dehydrogenase	264905, 264600, 264604, 264486
335	60057028 (669, 670)	Novel Protein sim. GBank gl 2103938 emb CAB096021 - (Z98800) glpQ2 [Mycobacterium tuberculosis]		esterase	264907, 264603, 264693, 18108374, 264636, 18108387
336	80414318 (671, 672)			UNCLASSIFIED	265009, 264766, 264686
337	11090829 (673, 674)				264602
338	95413134 (675, 676)	Novel Protein sim. GBank gl 5454074 ref NP_006303.1 pSMRT - silencing mediator for retinoid and thyroid hormone receptors	Contains protein domain (PF00249) - Myb-like DNA-binding domain	nuc_rept	264569, 18108397, 22278998, 29331822, 20281099, 29331824, 58182181, 68714117, 29331825, 35696052, 29331828, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 265008, 265009, 264910, 265009, 264758, 55812038, 65274444, 265011, 87188559, 265017, 265018, 265019, 264760, 55811150, 264681, 264762, 18108351, 264682, 264764, 264766, 264685, 264686, 264768, 52844229, 264689, 55811957, 35895917, 264692, 264693, 264628, 18108370, 18108374, 55811576, 35696423, 35695855, 264635, 264555, 264638, 264556, 264637, 264557, 18108360, 264638, 264558, 264639, 18108381, 83373044, 18108389, 87188518, 60432113
339	11398513 (677, 678)	Novel Protein sim. GBank gl 4001713 db BAA35087.1 - (AB015879) DnaK [Pseudomonas gingivalis]		eph	264593
340	80504149 (679, 680)	Novel Protein sim. GBank gl 2842699 sp Q92353 UBPC_SCHPO - PUTATIVE UBIQUITIN CARBOXYL-TERMINAL HYDROLASE C8G.08 (UBIQUITIN THIOLESTERASE) (UBIQUITIN-SPECIFIC PROCESSING PROTEASE) (DEUBIQUITINATING ENZYME)		ubiquitin	264905, 265019, 264769, 18108374
341	11075198 (681, 682)	Novel Protein sim. GBank gl 2688560 (AE001166) - conserved hypothetical protein [Borrelia burgdorferi]	Contains protein domain (PF00290) - Tryptophan synthase alpha chain	isomerase	264605
342	80054188 (683, 684)	Novel Protein sim. GBank gl 1684738 emb CAA70601 - (Y09452) YedJ hypothetical protein [Pseudomonas syringae]			264603, 264604
343	20466782 (685, 686)				264605
344	80428670 (687, 688)	Novel Protein sim. GBank gl 2117275 emb CAB091041 - (Z95818) hypothetical protein Rv0807 [Mycobacterium tuberculosis]		UNCLASSIFIED	264600, 264605, 264766, 18108370, 18108374, 35695855
345	80258853 (689, 690)	Novel Protein sim. GBank gl 3023317 sp Q48935 APHA_MYCRA - ACETYL POLYAMINE AMINOHYDROLASE		histone	264593
346	78831058 (691, 692)	Novel Protein sim. GBank gl 4239787 emb CAA75437 - (Y15166) NADP-glutamate dehydrogenase [Pseudomonas aeruginosa]	Contains protein domain (PF00208) - Glutamate/Leucine/Phenylalanine/Va line dehydrogenase	dehydrogenase	264905

347	79158185 (693, 694)	Novel Protein sim. GBank gi 731675 sp P38795 YHM4_YEAST - HYPOTHETICAL 80.7 KD PROTEIN IN ERG7-NMD2 INTERGENIC REGION		UNCLASSIFIED	265006, 265008, 265010, 265018, 263987, 263981
348	80020208 (695, 698)	Novel Protein sim. GBank gi 1073610 sp S47672 - ugpB protein - Escherichia coli		transport	264602, 18108351, 18108387
349	17282112 (697, 698)	Novel Protein sim. GBank gi 3261599 emb CAB009171 - (Z77137) hypothetical protein Rv1277 [Mycobacterium tuberculosis]		nuclease	265007
350	80502370 (699, 700)	Novel Protein sim. GBank gi 2959367 emb CAA178211 - (AL022117) hypothetical protein [Schizosaccharomyces pombe]			265009, 264769, 264689, 18108370
351	80501805 (701, 702)	Novel Protein sim. GBank gi 4416302 gb AAD203071 - (AF105716) copia-type pol polyprotein [Zea mays]		glycoprotein	264769, 264905, 264908
352	11611585 (703, 704)	Novel Protein sim. GBank gi 174887 sp P42873 URE1_STAXY - UREASE ALPHA SUBUNIT (UREA AMIDOHYDROLASE)	Contains protein domain (PF00449) - Urease	protease	264595
353	80061653 (705, 706)	Novel Protein sim. GBank gi 115157 sp P16574 BVGA_BORPE - VIRULENCE FACTORS PUTATIVE POSITIVE TRANSCRIPTION REGULATOR BVGA		UNCLASSIFIED	264604
354	56626130 (707, 708)	Novel Protein sim. GBank gi 115157 sp P16574 BVGA_BORPE - VIRULENCE FACTORS PUTATIVE POSITIVE TRANSCRIPTION REGULATOR BVGA	Contains protein domain (PF00072) - Response regulator receiver domain	UNCLASSIFIED	264628
355	80046344 (709, 710)	Novel Protein sim. GBank gi 115157 sp P16574 BVGA_BORPE - VIRULENCE FACTORS PUTATIVE POSITIVE TRANSCRIPTION REGULATOR BVGA		UNCLASSIFIED	264909, 264595, 264683, 22279002
356	80043835 (711, 712)	Novel Protein sim. GBank gi 115157 sp P16574 BVGA_BORPE - VIRULENCE FACTORS PUTATIVE POSITIVE TRANSCRIPTION REGULATOR BVGA		transcript factor	264909, 264591, 264592
357	80070568 (713, 714)	Novel Protein sim. GBank gi 497637 (J03839) - cytochrome oxidase d subunit I [Escherichia coli]		oxidase	264605
358	37032756 (715, 716)	Novel Protein sim. GBank gi 2280990 (AF006000) - Btg1 [Bordetella pertussis]		UNCLASSIFIED	264768
359	80501488 (717, 718)	Novel Protein sim. GBank gi 3510639 (AF049344) - UDP- GalNAc:polypeptide N-acetylglucosaminyltransferase T5 [Rattus norvegicus]		UNCLASSIFIED	264604, 264769
360	80026748 (719, 720)	Novel Protein sim. GBank gi 113764 sp P25718 AMY1 - ALPHA-AMYLASE PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE)		UNCLASSIFIED	264594
361	80584075 (721, 722)	Novel Protein sim. GBank gi 113764 sp P25718 AMY1 - ALPHA-AMYLASE PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE)		transferase	22278996, 264259, 28331822, 29331824, 264605, 55811957, 265022
362	13089485 (723, 724)	Novel Protein sim. GBank gi 113764 sp P25718 AMY1 - ALPHA-AMYLASE PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE)		amylase	264688
363	79750145 (725, 726)	Novel Protein sim. GBank gi 2029818 sp P95171 NUOK_MYCTU - NADH DEHYDROGENASE I CHAIN K (NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 1) (NUO11)	Contains protein domain (PF00420) - NADH-ubiquinone/plastoquinone oxidoreductase chain 4L		264568
364	82443593 (727, 728)	Novel Protein sim. GBank gi 2029818 sp P95171 NUOK_MYCTU - NADH DEHYDROGENASE I CHAIN K (NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 1) (NUO11)		dehydrogenase	264769, 264602, 264604, 264508, 264762, 264638, 264486

365	88040288 (728, 730)	Novel Protein sim. GBank glj4926268[gbjAAD33924.1] - (AF144237) LOMP protein [Homo sapiens]	Contains protein domain (PF00412) - LIM domain containing proteins		284488, 21808768, 21808787, 55811578, 21908769, 29148829, 22278985, 22278986, 285020, 265022, 264634, 264691, 284593, 33857023, 33857402, 264693, 264639, 284594, 29331824, 264758, 18108385, 29331827, 87168559, 265018, 22278000, 265019, 264482, 264761, 264681, 18108351, 265017, 264757
366	81821838 (731, 732)				
367	95357471 (733, 734)	Novel Protein sim. GBank glj4503843[re]NP_003908.1pG2AD - UNKNOWN	Contains protein domain (PF01602) - Apatin N terminal region	glycoprotein	60424178, 65274572, 58182575, 22278994, 58994075, 22278998, 264259, 29331822, 29331824, 58182181, 60424289, 88714117, 29331825, 60432289, 29331826, 29331827, 29331828, 284905, 284828, 58182435, 265006, 264512, 265008, 264591, 55812038, 55811386, 265010, 87188559, 265017, 265018, 264604, 265019, 55811150, 264448, 264369, 264288, 264886, 264768, 58181562, 21808768, 21808769, 55811957, 35895917, 265022, 60170815, 33857023, 85274620, 18108385, 263967, 33857109, 33857349, 35895763, 264628, 18108376, 55811578, 85274791, 35895855, 58182323, 83373044, 80432113, 264583, 264584, 284587, 264509
368	79807265 (735, 736)	Novel Protein sim. GBank glj3913029[sp]P94967[ALR_MYCSM - ALANINE RACEMASE		UNCLASSIFIED	264508, 264604, 264605, 264636
369	95292917 (737, 738)				
370	88090968 (739, 740)	Novel Protein sim. GBank glj3249559 (AF018261) - EH domain binding protein Epsin [Rattus norvegicus]			284905, 264592, 284605, 264766, 264691
371	95292589 (741, 742)	Novel Protein sim. GBank glj2995299[emb]CAA18328] - (AL022268) putative tRNA delta(2)- isopentenylpyrophosphate transferase [Streptomyces coelicolor]	Contains protein domain (PF01715) - IPP transferase	transferase	264905, 284906, 284510, 264800, 264601, 264602, 264603, 265018, 264604, 264605, 265021, 284692, 264636, 264564
372	80021107 (743, 744)	Novel Protein sim. GBank glj2508393[ep]P31576[FIXX_ECOLI - FERREDOXIN LIKE PROTEIN			264564
373	79863768 (745, 746)				
374	79847568 (747, 748)	Novel Protein sim. GBank glj334164[emb]CAA13164] - (AJ231122) z81f [Vibrio cholerae]		UNCLASSIFIED	264909
375	91230181 (749, 750)	Novel Protein sim. GBank glj5456834[gbjAAD43716.1] - (AF152322) prolactin gamma A2 [Homo sapiens]		UNCLASSIFIED	284905, 264906
376	80505214 (751, 752)	Novel Protein sim. GBank glj1805409[dbj]BAA08970] - (D50453) homologues to nitrite hydratase region 3' - hypothetical protein P47K of P. chlororaphis [Bacillus subtilis]		cadherin	65274572, 284259, 28331828, 58182435, 60433356, 60433438, 284757, 55812038, 264758, 55811857, 264690, 33657023, 264769
377	10339083 (753, 754)			UNCLASSIFIED	264769
					264906

378	80056153 (755, 756)	Novel Protein sim. GBank gij1076013[pir][A49930 - carB (fragment)]	Contains protein domain (PF00289) - Carbamoyl-phosphate synthase (CPSase)	UNCLASSIFIED	265008, 264555
379	80503437 (757, 758)	protein homolog - <i>Mycobacterium bovis</i> (strain BCG)			264769
380	80060937 (759, 760)	Novel Protein sim. GBank gij216556[idb][BAA02174] - (D12851) glucose dehydrogenase [<i>Escherichia coli</i>]	Contains protein domain (PF01011) - PQQ enzyme repeat	dehydrogenase	264604
381	11769027 (761, 762)			UNCLASSIFIED	264684
382	80054377 (763, 764)				264592
383	83259025 (765, 766)	Novel Protein sim. GBank gij3327136[idb][BAA31636] - (AB014581) KIAA0661 protein [<i>Homo sapiens</i>]			264595, 265017, 265021, 264638, 67168518, 22279002
384	95314255 (767, 768)			UNCLASSIFIED	264259, 29331822, 60432289, 29331827, 264288, 264768, 263987, 85274791, 35695855, 263981, 83373044, 264567
385	10237879 (769, 770)	Novel Protein sim. GBank gij1073456[pir][S47810 - probable alcohol dehydrogenase (EC 1.1.1.1) - <i>Escherichia coli</i>]	Contains protein domain (PF00465) - Iron-containing alcohol dehydrogenases	dehydrogenase	264692
386	76633434 (771, 772)				264908
387	17860637 (773, 774)	Novel Protein sim. GBank gij1460074[emb][CAB01049] - (Z77250) hypothetical protein Rv2566 [<i>Mycobacterium tuberculosis</i>]	Contains protein domain (PF01841) - Transglutaminase-like superfamily	UNCLASSIFIED	264760
388	87741376 (775, 776)	Novel Protein sim. GBank gij4240169[idb][BAA74863.1] - (AB020847) KIAA0840 protein [<i>Homo sapiens</i>]	Contains protein domain (PF00848) - F-box domain.	homeobox	35696286, 264905, 68712502, 60432229, 264593, 60433358, 264688, 264688, 21908765, 264891, 22279000, 264482
389	79316971 (777, 778)			UNCLASSIFIED	18108394, 22278986, 264630, 264558, 22279002
390	80079949 (779, 780)	Novel Protein sim. GBank gij654065[emb][CAA58337] - (X83413) U88 [<i>Human herpesvirus 6</i>]		UNCLASSIFIED	264600
391	7657302 (781, 782)				264482
392	79796056 (783, 784)	Novel Protein sim. GBank gij3378523[emb][CAA08867] - (AJ009832) cyclomaltoextrinase glucanotransferase [<i>Thermoplasma neapolitana</i>]		UNCLASSIFIED	264908
393	33206031 (785, 786)			synthase	264602, 21908764
394	10104463 (787, 788)				264693
395	80229010 (789, 790)	Novel Protein sim. GBank gij2677780 (U70327) - unknown		UNCLASSIFIED	264508, 264563
396	20436224 (791, 792)	[<i>Paretioplus polyacis</i>]	Contains protein domain (PF00047) - Immunoglobulin domain	strut	264556
397	80417014 (793, 794)	Novel Protein sim. GBank gij4507909[ref]NP_000368.1[pWAS] - Wiskott-Aldrich syndrome (eczema-thrombocytopenia)			265007, 265009, 264508, 264556, 264629, 264766
398	91230517 (795, 796)	Novel Protein sim. GBank gij1518458 (U45998) - mitochondrial solute carrier [<i>Onchocerca volvulus</i>]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	transport	18108398, 22278995, 22278996, 56994075, 22278999, 264259, 29331824, 29331828, 264905, 264908, 265007, 265008, 265009, 21908754, 33657084, 265017, 264448, 264288, 264768, 21908785, 21908768, 21908767, 265020, 265021, 33657023, 33657109, 264628, 35698423, 35695855, 264952, 18108380, 264587, 18108391

399	80055278 (797, 798)	Novel Protein sim. GBank gij3358091[dbj BAA31895] - (AB015974) glycerol kinase [Pseudomonas tolaasii]	Contains protein domain (PF00370) - kinases FGGY family of carbohydrate	kinase	284592, 284595
400	94117490 (799, 800)	Novel Protein sim. GBank gij728835[sp P39192 ALU5_HUMAN - IIII ALU SUBFAMILY SC WARNING ENTRY IIII]	Contains protein domain (PF00580) - Leucine Rich Repeat	cadherin	18108394, 56182575, 22278995, 22278997, 22278999, 284259, 29331824, 285008, 285007, 285009, 60432229, 33657402, 21908754, 285010, 285017, 285018, 285019, 18108351, 18108357, 21908765, 285021, 285022, 284681, 284692, 33657023, 18108370, 65274791, 284634, 284636, 60170394, 56182323
401	11397491 (801, 802)	Novel Protein sim. GBank gij4828292[gb AAD33527.1 AF132117] FhuA [Staphylococcus aureus]		transport	284594
402	95420294 (803, 804)	Novel Protein sim. GBank gij5689487[dbj BAA83027.1] - (AB028998) KIAA1075 protein [Homo sapiens]	Contains protein domain (PF00017) - Src homology domain 2	phosphatase	65274572, 56182575, 35898286, 22278996, 22278998, 284093, 284259, 29331822, 29331824, 29331825, 29331826, 60432289, 29331827, 29331828, 284908, 284907, 284909, 285006, 284511, 285007, 285008, 284910, 284591, 33657402, 60433356, 60433438, 284598, 21908754, 52844286, 265010, 285011, 87168559, 265017, 265018, 265019, 284681, 18108351, 284682, 284448, 284288, 284684, 284766, 264767, 264688, 21908765, 21908768, 21908767, 21908768, 21908769, 55811957, 285020, 285021, 285022, 284690, 284693, 85274620, 35695763, 284628, 18108370, 284629, 18108378, 35898423, 55811576, 284635, 284636, 284557, 284639, 18108385, 22279002, 284563, 284564, 284585, 284568, 284768, 284632, 284639, 284563
403	80439913 (805, 806)			UNCLASSIFIED	284682
404	11809885 (807, 808)			polymerase	285009, 284682
405	79471280 (809, 810)	Novel Protein sim. GBank gij2861649[emb CAA15755] - (AL009198) dnaE2 [Mycobacterium tuberculosis]		UNCLASSIFIED	18108357, 284693
406	79634172 (811, 812)		Contains protein domain (PF00159) - Pancreatic hormone peptides	UNCLASSIFIED	
407	80478229 (813, 814)			UNCLASSIFIED	284769
408	80078958 (815, 816)			UNCLASSIFIED	284600
409	5840527 (817, 818)	Novel Protein sim. GBank gij3047117 (AF058919) - similar to ATP-dependent RNA helicases [Arabidopsis thaliana]		helicase	284259

410	95357486 (818, 820)	Novel Protein sim. GBank gij475016[dj]BAA08184] - (D29801) Unknown [Mus musculus]		UNCLASSIFIED	264489, 52846365, 52846842, 56181886, 35896286, 52845080, 29331822, 29331824, 56182181, 29331825, 60424289, 35898052, 33656970, 264508, 264509, 264905, 264906, 264907, 264908, 52844045, 264909, 264510, 265007, 264512, 265008, 264910, 33657402, 284758, 52846317, 55811388, 285010, 265011, 265017, 264604, 285018, 55811150, 264782, 264764, 264786, 264887, 264788, 264769, 52844229, 21906786, 265020, 285021, 264534, 52844150, 264892, 33657023, 65274820, 33657109, 33657182, 27488281, 35895783, 264828, 264829, 60431528, 18108376, 263978, 35898423, 35895855, 264832, 264834, 264835, 264637, 264838, 264558, 264839, 56182323, 284559, 60432113, 22279002, 284563, 284565, 284486
411	80501670 (821, 822)			UNCLASSIFIED	264789
412	80241662 (823, 824)				284807, 264910, 263973, 22278002
413	11078446 (825, 826)	Novel Protein sim. GBank gij3281784[emb]CAB08897] - (Z95558) htpX [Mycobacterium tuberculosis]		eph	264605
414	82050554 (827, 828)	Novel Protein sim. GBank gij129038[sp]P20707[DDO1_AZOV] - 2-OXOGLUTARATE DEHYDROGENASE E1 COMPONENT (ALPHA-KETOGLUTARATE DEHYDROGENASE)		dehydrogenase	18108374, 264760, 264769, 264602, 284638, 284603, 264908, 264805
415	84453144 (828, 830)	Novel Protein sim. GBank gij4868350[gb]AAD31273.1[AF13202 - (AF132025) rhophilin [Drosophila melanogaster]		UNCLASSIFIED	264908, 87188518
416	60402775 (831, 832)	Novel Protein sim. GBank gij2555172 (AF025543) - ArcC; carbamate kinase [Rhizobium elii]		kinase	284488, 264600, 264602, 264764, 264636
417	20153787 (833, 834)	Novel Protein sim. GBank gij170917[sp]P52311[MTX2_XANOR - MODIFICATION METHYLASE XORII (CYTOSINE-SPECIFIC METHYLTRANSFERASE XORII) (M.XORII)]	Contains protein domain (PF00145) - C-5 cytosine-specific DNA methylase		264605
418	94125841 (835, 836)			UNCLASSIFIED	284889, 264893
419	95314273 (837, 838)			collagen	264908, 264910, 264764, 264839
420	37038349 (839, 840)	Novel Protein sim. GBank gij3281659[emb]CAB03751] - (Z81368) hypothetical protein Rv2418c [Mycobacterium tuberculosis]	Contains protein domain (PF00300) - Phosphoglycerate mutase family	phosphatase	264789
421	95292942 (841, 842)	Novel Protein sim. GBank gij2916942[emb]CAA17580] - (AL021999) hypothetical protein Rv0981 [Mycobacterium tuberculosis]	Contains protein domain (PF00072) - Response regulator receiver domain	phosphatase	264908, 264600, 264801, 264603, 264604, 264760, 264789
422	78471293 (843, 844)	Novel Protein sim. GBank gij231752[sp]Q00767[CH81_STRAL - 80 KD CHAPERONIN 1 (PROTEIN CPN60 1) (GROEL PROTEIN 1) (HSP58)]	Contains protein domain (PF00118) - TCP-1/cpn60 chaperonin family	eph	22278996, 264682, 18108376, 18108387
423	79604948 (845, 846)			UNCLASSIFIED	264509

424	78966557 (847, 846)	Novel Protein sim. GBank gi 4826814 ref NP_004977.1 pKTN1 - kinesin receptor			struct	265019	
425	80431450 (849, 850)	Novel Protein sim. GBank gi 1703701 bbs 170462 - KRPS=kinesin-related protein [rats, testes, Peptide Partial, 187 aa]		Contains protein domain (PF002225) - Kinesin motor domain	struct	264909, 265007, 55811386, 264766, 55810764	
426	80064522 (851, 852)	Novel Protein sim. GBank			UNCLASSIFIED	264605, 264559	
427	80057232 (853, 854)	Novel Protein sim. GBank gi 231829 sp P29929 COBN_PSEDE - COBN PROTEIN			UNCLASSIFIED	264603, 264636	
428	79487798 (855, 856)	Novel Protein sim. GBank gi 61286 p S22697 - extensin - Vohox cartier (fragment)			UNCLASSIFIED	264683	
429	80091252 (857, 858)	Novel Protein sim. GBank gi 1808154 emb CAB06451 - (Z84395) hypothetical protein Rv0688 [Mycobacterium tuberculosis]			UNCLASSIFIED	35698423, 35695763, 35695655, 265017, 264584, 264762	
430	80504192 (859, 860)	Novel Protein sim. GBank gi 1808154 emb CAB06451 - (Z84395) hypothetical protein Rv0688 [Mycobacterium tuberculosis]			reductase	264508, 264805, 264509, 264908, 264809, 265008, 264600, 264887, 264789, 264689, 264636, 264638, 18108385, 264486	
431	20824248 (861, 862)	Novel Protein sim. GBank gi 3123552 emb CAA18609 - (AL022578) dJ393P12.2 (hypothetical Proline-rich protein KIAA0269 LIKE) [Homo sapiens]			UNCLASSIFIED	264568	
432	16525372 (863, 864)	Novel Protein sim. GBank gi 12495272 sp Q89826 CDX2_HUMAN - HOMEBOX PROTEIN CDX-2 (CAUDAL-TYPE HOMEBOX PROTEIN 2) (CDX-3)			UNCLASSIFIED	265020	
433	81494303 (865, 866)	Novel Protein sim. GBank gi 114105 sp P08532 ARAH_ECOLI - L-ARABINOSE TRANSPORT SYSTEM PERMEASE PROTEIN ARAH			UNCLASSIFIED	264907, 264908, 264909, 264910, 264592, 264595, 264758, 264804, 264760, 264762, 264763, 264636, 264637, 22279002	
434	94326323 (867, 868)	Novel Protein sim. GBank gi 2495272 sp Q89826 CDX2_HUMAN - HOMEBOX PROTEIN CDX-2 (CAUDAL-TYPE HOMEBOX PROTEIN 2) (CDX-3)		Contains protein domain (PF00169) - PH domain	UNCLASSIFIED	55812038, 58182181, 58181562, 28331828, 35696052, 55810764, 55811578, 65274791, 35695855, 60432113, 55811150, 264636, 264766	
435	80502738 (869, 870)	Novel Protein sim. GBank gi 114105 sp P08532 ARAH_ECOLI - L-ARABINOSE TRANSPORT SYSTEM PERMEASE PROTEIN ARAH			transport	264595, 264769	
436	41085953 (871, 872)	Novel Protein sim. GBank gi 3242702 (AC003040) - hypothetical protein [Arabidopsis thaliana]			UNCLASSIFIED	265020, 22279002	
437	1139291 (873, 874)	Novel Protein sim. GBank gi 2291232 gb AA085351.1 - (AF018427) Contains similarity to Pfam domain: PF00004 (AAA). Score=288.1, E-value=3.7e-77, N=1 [Caenorhabditis elegans]		Contains protein domain (PF00004) - ATPases associated with various cellular activities (AAA)	UNCLASSIFIED	264593	
438	11773635 (875, 876)	Novel Protein sim. GBank gi 3242702 (AC003040) - hypothetical protein [Arabidopsis thaliana]			UNCLASSIFIED	264686	
439	80018495 (877, 878)	Novel Protein sim. GBank gi 3242702 (AC003040) - hypothetical protein [Arabidopsis thaliana]				264905, 264600, 264602, 264604	
440	79841062 (879, 880)	Novel Protein sim. GBank gi 2291232 gb AA085351.1 - (AF018427) Contains similarity to Pfam domain: PF00004 (AAA). Score=288.1, E-value=3.7e-77, N=1 [Caenorhabditis elegans]		Contains protein domain (PF00004) - ATPases associated with various cellular activities (AAA)	ATPase, associated	35698052, 264905, 264908, 264909, 265011, 35698423	
441	20396935 (881, 882)	Novel Protein sim. GBank gi 5639946 gb AAD45904.1 AF16132 - (AF161328) histidine kinase CstS [Corynebacterium diptheriae]				264605	
442	85281058 (883, 884)	Novel Protein sim. GBank gi 1184780 (U46088) - von Ebner minor salivary gland protein [Mus musculus]			UNCLASSIFIED	29331830, 264909	
443	82458427 (885, 886)	Novel Protein sim. GBank gi 5689893 emb CAB52058.1 - (AL109732) putative ATP-binding RNA helicase [Streptomyces coelicolor A3(2)]			UNCLASSIFIED	35696052, 264508, 264906, 264512, 264604, 264762, 264769, 264889, 264636	
444	11395897 (887, 888)	Novel Protein sim. GBank gi 1783249 dbj BAA11726 - (D83026) homologous to citrate-sodium symport (citrate transporters), hypothetical [Bacillus subtilis]			UNCLASSIFIED	264591	

445	70552709 (889, 890)	Novel Protein sim. GBank gl 5531272[emb CAB50897.1] - (AJ243800) WSC4 homologue [Kluyveromyces fragilis]		UNCLASSIFIED	264693
446	79810937 (891, 892)	Novel Protein sim. GBank gl 538413 (L36315) - zinc finger protein [Mus musculus]	Contains protein domain (PF00099) - Zinc finger, C2H2 type		264509
447	80438888 (893, 894)	Novel Protein sim. GBank gl 1542914[emb CAB02185] - (Z80108) fml [Mycobacterium tuberculosis]	Contains protein domain (PF00551) - Formyl transferase	transcript factor	264768, 55811576
448	80238110 (895, 896)	Novel Protein sim. GBank gl 118794[sp P10443]DP3A_ECOLI - DNA POLYMERASE III, ALPHA CHAIN		dehydrogenase	264508, 264600, 264603, 264605, 264682, 264769, 18108382, 284634, 18108387
449	20460834 (897, 898)	Novel Protein sim. GBank gl 4589506[db BAA76775.1] - (AB023148) KIAA0931 protein [Homo sapiens]		polymerase	264605, 264559
450	84631210 (899, 900)	Novel Protein sim. GBank gl 2650614 (AE001104) - conserved hypothetical protein [Archaeoglobus fulgidus]	Contains protein domain (PF00481) - Protein phosphatase 2C		65274572, 22278898, 28331824, 28331826, 284906, 264910, 284592, 52646317, 265017, 21908767, 55811957, 56528486, 22279002
451	21433609 (901, 902)	Novel Protein sim. GBank gl 2493000[sp Q09450]SCOT_CAEEL - PROBABLE SUCCINYL-COA:3-KETOACID-COENZYME A TRANSFERASE PRECURSOR (3-OXOACID COA-TRANSFERASE)		UNCLASSIFIED	264486
452	10287278 (903, 904)	Novel Protein sim. GBank gl 3411053 (AF034863) - synaptic scaffolding molecule [Rattus norvegicus]		UNCLASSIFIED	284692
453	52560098 (905, 906)	Novel Protein sim. GBank gl 4468699[emb CAB38153.1] - (AL035591) putative integral membrane export protein [Streptomyces coelicolor]		UNCLASSIFIED	284691
454	39523922 (907, 908)	Novel Protein sim. GBank gl 4589506[db BAA76775.1] - (AB023148) KIAA0931 protein [Homo sapiens]		UNCLASSIFIED	284905
455	13089692 (909, 910)	Novel Protein sim. GBank gl 4589506[db BAA76775.1] - (AB023148) KIAA0931 protein [Homo sapiens]		UNCLASSIFIED	284687
456	79583081 (911, 912)	Novel Protein sim. GBank gl 4589506[db BAA76775.1] - (AB023148) KIAA0931 protein [Homo sapiens]		UNCLASSIFIED	284691
457	79831273 (913, 914)	Novel Protein sim. GBank gl 4589506[db BAA76775.1] - (AB023148) KIAA0931 protein [Homo sapiens]		UNCLASSIFIED	284905
458	79581227 (915, 916)	Novel Protein sim. GBank gl 4589506[db BAA76775.1] - (AB023148) KIAA0931 protein [Homo sapiens]		UNCLASSIFIED	284687
459	80567359 (917, 918)	Novel Protein sim. GBank gl 4589506[db BAA76775.1] - (AB023148) KIAA0931 protein [Homo sapiens]		UNCLASSIFIED	284687
460	79245890 (919, 920)	Novel Protein sim. GBank gl 4589506[db BAA76775.1] - (AB023148) KIAA0931 protein [Homo sapiens]		UNCLASSIFIED	284687
461	95287618 (921, 922)	Novel Protein sim. GBank gl 4589506[db BAA76775.1] - (AB023148) KIAA0931 protein [Homo sapiens]		UNCLASSIFIED	284687

462	79806588 (923, 924)	Novel Protein sim. GBank gi1346891 pp 45597 TF1_XANCP - MULTIPHOSPHORYL TRANSFER PROTEIN (MTP) - (CONTAINS: PHOSPHOENOLPYRUVATE-PROTEIN PHOSPHOTRANSFERASE (PHOSPHOTRANSFERASE SYSTEM, ENZYME I); PHOSPHOCARRIER PROTEIN HPR (PROTEIN H); PTS SYSTEM, FRUCTOSE-SPECIFIC IIA COMPONENT ... (X83413) U88 (Human herpesvirus 6)	Contains protein domain (PF00391) - PEP-utilizing enzymes	UNCLASSIFIED	284807
463	79786417 (925, 926)	Novel Protein sim. GBank gi1854065 emb CAA583371 - (X83413) U88 (Human herpesvirus 6)		UNCLASSIFIED	284805, 284806, 284808, 284809, 284910, 284591, 284595, 285011, 284632, 284635, 284638, 284637, 284638, 284639
464	82340151 (927, 928)	Novel Protein sim. GBank gi15889778 emb CAB52137.1 - (AJ242832) calpain [Homo sapiens]	Contains protein domain (PF00848) - Calpain family cysteine protease	UNCLASSIFIED	284834
465	83005730 (929, 930)	Novel Protein sim. GBank gi1808175 emb CAB06470 - (Z84395) rpsC [Mycobacterium tuberculosis]	Contains protein domain (PF00417) - Ribosomal protein S3, N-terminal domain.	calthepsin ribosomal prot	285017, 21906784, 265020 284605, 284559
466	20460645 (931, 932)	Novel Protein sim. GBank gi1543705 pp 36849 RBSB_BACSU - D-RIBOSE-BINDING PROTEIN PRECURSOR		UNCLASSIFIED	264784
467	60408035 (933, 934)	Novel Protein sim. GBank gi2114024 emb CAB089571 - (Z95559) grcC1 [Mycobacterium tuberculosis]		UNCLASSIFIED	284692
468	52562208 (935, 936)	Novel Protein sim. GBank gi2809459 emb CAA173471 - (AL021929) cobQ [Mycobacterium tuberculosis]		UNCLASSIFIED	284488
469	18520527 (937, 938)	Novel Protein sim. GBank gi114921 pp 17447 BETT_ECOLI - HIGH-AFFINITY CHOLINE TRANSPORT PROTEIN		synthase transport	284602, 284769 265019
470	80502758 (939, 940)	Novel Protein sim. GBank gi114921 pp 17447 BETT_ECOLI - HIGH-AFFINITY CHOLINE TRANSPORT PROTEIN			
471	17937351 (941, 942)	Novel Protein sim. GBank gi114921 pp 17447 BETT_ECOLI - HIGH-AFFINITY CHOLINE TRANSPORT PROTEIN			
472	80047458 (943, 944)	Novel Protein sim. GBank gi1862343 (L10808) - Gcap1 gene product [Mus musculus]		UNCLASSIFIED	284598, 284685, 284557 284369
473	20558783 (945, 946)	Novel Protein sim. GBank gi1862343 (L10808) - Gcap1 gene product [Mus musculus]		UNCLASSIFIED	22278997, 284692, 284288 284907, 284908, 284511, 265009, 284782, 284448, 284638, 284638
474	80593365 (947, 948)	Novel Protein sim. GBank gi15453656 jnp_006329.1 pGAC1 - glioma amplified on chromosome 1 protein (leucine-rich)	Contains protein domain (PF00560) - Leucine Rich Repeat	glycoprotein	65274572, 60432049, 284259, 284508, 52644045, 55812038, 284758, 265011, 284288, 284688, 52844229, 65274791, 284638, 284586
475	82454685 (949, 950)	Novel Protein sim. GBank gi15453656 jnp_006329.1 pGAC1 - glioma amplified on chromosome 1 protein (leucine-rich)		UNCLASSIFIED	284638
476	94143857 (951, 952)	Novel Protein sim. GBank gi15453656 jnp_006329.1 pGAC1 - glioma amplified on chromosome 1 protein (leucine-rich)		UNCLASSIFIED	284690, 284693
477	79175833 (953, 954)	Novel Protein sim. GBank gi15453656 jnp_006329.1 pGAC1 - glioma amplified on chromosome 1 protein (leucine-rich)		UNCLASSIFIED	284686, 35695855, 285008, 284631, 284910, 284632, 284638, 265018, 284369
478	78633463 (955, 956)	Novel Protein sim. GBank gi1127551 (U18939) - orf2 [Batrachocottus baikalensis]		collagen	284369
479	80189746 (957, 958)	Novel Protein sim. GBank gi1127551 (U18939) - orf2 [Batrachocottus baikalensis]		mapolymerase	
480	79390729 (959, 960)	Novel Protein sim. GBank gi1127551 (U18939) - orf2 [Batrachocottus baikalensis]			
481	78624578 (961, 962)	Novel Protein sim. GBank gi14083042 (AF068065) - GP900; mucin-like glycoprotein [Cryptosporidium parvum]		UNCLASSIFIED	284693
482	83050611 (963, 964)	Novel Protein sim. GBank gi14083042 (AF068065) - GP900; mucin-like glycoprotein [Cryptosporidium parvum]		UNCLASSIFIED	284809, 284868, 284768, 284693, 55811576, 56182323, 18108385

483	20293306 (965, 966)	Novel Protein sim. GBank gl 2104303 emb CA080632 - (Z95387) hypothetical protein Rv2610c [Mycobacterium tuberculosis]	Contains protein domain (PF00534) - Glycosyl transferases group 1	284600	
484	11618046 (887, 968)	Novel Protein sim. GBank gl 3450883 (AF083334) - fibrin [Anthraxa permy]	UNCLASSIFIED	284594	
485	80191234 (989, 970)	Novel Protein sim. GBank gl 5042272 emb CAB44526.1 - (AL078618) nuoF, NADH dehydrogenase subunit [Streptomyces coelicolor]	UNCLASSIFIED	284369, 21808785, 22278000, 22278002	
486	80059042 (871, 972)		dehydrogenase	284604	
487	11813339 (973, 974)			284638	
488	91222383 (975, 976)	Novel Protein sim. GBank gl 5724778 gb AAC53522.2 - (AF012273) rho-type GTPase-activating protein rhoGAPX-1 [Mus musculus]	Contains protein domain (PF00620) - RhoGAP domain	284686, 66714117, 284768, 18108385, 55811576, 265006, 285008, 265009, 265019, 22278002, 264259, 18108370, 264907, 284784, 58182323, 284288, 284693	
489	10867710 (977, 978)	Novel Protein sim. GBank gl 3882223 dbj BAA34471.1 - (AB018294) KIAA0751 protein [Homo sapiens]	Kinase	284639	
490	95361124 (878, 980)	Novel Protein sim. GBank gl 82091 pir A25494 - hydroxyproline-rich glycoprotein - tomato (fragment)	collagen	22278996, 29331822, 29331828, 284107, 284908, 264110, 265009, 264592, 284593, 60433358, 284288, 284683, 263974, 263976, 20281071, 60432113	
491	80488412 (981, 982)	Novel Protein sim. GBank gl 2894206 emb CAA17072 - (AL021840) hypothetical protein Rv3258c [Mycobacterium tuberculosis]	UNCLASSIFIED	284769	
492	87421264 (983, 984)			284600	
493	11692842 (985, 986)		UNCLASSIFIED	284638	
494	87726604 (987, 988)	Novel Protein sim. GBank gl 5262605 emb CAB45743.1 - (AL080150) hypothetical protein [Homo sapiens]	UNCLASSIFIED	284489, 35696288, 60432289, 29331828, 35698052, 284509, 284905, 284908, 284907, 284908, 264909, 284510, 284511, 265009, 284910, 33657402, 284762, 284764, 284768, 284769, 284888, 21906785, 21906769, 35695917, 285020, 284693, 33657108, 284629, 35696423, 35695855, 284634, 284638	
495	80028599 (989, 990)	Novel Protein sim. GBank gl 2791517 emb CAA16054 - (AL021248) hypothetical protein Rv2477c [Mycobacterium tuberculosis]	Contains protein domain (PF00005) - ABC transporter	284602, 264682, 264638	
496	78985624 (991, 992)	Novel Protein sim. GBank gl 230281 pdb 1R69 - 434 Repressor (Amino-Terminal Domain) (R1-69)		284601, 265021	
497	78948661 (993, 994)	Novel Protein sim. GBank gl 128738 sp P28225 PDXH_ECOLI - PYRIDOXAMINE 5'-PHOSPHATE OXIDASE (PNP/IMP OXIDASE)	oxidase	265006	

488	88095488 (995, 998)	Novel Protein sim. GBank gij1145789 (U41662) - neuroligin 2 [Rattus norvegicus]	Contains protein domain (PF00135) - esterase Carboxylesterases	264259, 29331826, 35698052, 264508, 284509, 284905, 264908, 284907, 284908, 284909, 284510, 284511, 265009, 284910, 284591, 33857402, 284756, 265010, 265011, 284800, 284601, 284805, 264683, 264764, 284788, 284787, 284768, 284687, 284769, 21906787, 33857023, 284683, 284628, 284628, 35698423, 284830, 284632, 284834, 284635, 284637, 284638, 264558, 284639, 10108385, 284563, 284564, 284565, 284566, 284567
499	20438222 (987, 998)	Novel Protein sim. GBank gij97480[pil]S18739 - Integral membrane protein - Rhodobacter capsulatus	UNCLASSIFIED	284605
500	11076810 (999, 1000)			
501	13418034 (1001, 1002)	Novel Protein sim. GBank gij5708250[emb]CAB52363.1] - (AL108747) putative Integral membrane protein [Streptomyces coelicolor A3(2)]	UNCLASSIFIED	284688
502	80021176 (1003, 1004)	Novel Protein sim. GBank gij468678[emb]CAB38132.1] - (AL035591) glucose-6-phosphate isomerase [Streptomyces coelicolor]	Contains protein domain (PF00342) - isomerase Phosphoglucose isomerase	22278996, 265011, 284602, 284603, 284635
503	20284483 (1005, 1006)		UNCLASSIFIED	284584
504	10887321 (1007, 1008)		UNCLASSIFIED	284687
505	95003068 (1009, 1010)		UNCLASSIFIED	284389
506	16454282 (1011, 1012)		Contains protein domain (PF00036) - EF hand	265010
507	20451598 (1013, 1014)	Novel Protein sim. GBank gij4033509[sp]P02598[CALM_TETPY - CALMODULIN Novel Protein sim. GBank gij2501069[sp]Q46127[SYW_CLOLO - TRYPTOPHANYL-TRNA SYNTHETASE (TRYPTOPHAN--TRNA LIGASE) (TRPRS)	UNCLASSIFIED	264604
508	78841424 (1015, 1016)	Novel Protein sim. GBank gij466088[sp]P34618[YO82_CAEEL - HYPOTHETICAL 33.8 KD PROTEIN ZK1236.2 IN CHROMOSOME III	UNCLASSIFIED	284908
509	11776388 (1017, 1018)			284638
510	83373465 (1019, 1020)		UNCLASSIFIED	284687, 284639
511	16325578 (1021, 1022)			265007
512	20399484 (1023, 1024)	Novel Protein sim. GBank gij2497419[sp]P55635[Y4RB_RHISN - PUTATIVE INTEGRASE/RECOMBINASE Y4RB	UNCLASSIFIED	284585
513	79457404 (1025, 1026)	Novel Protein sim. GBank gij1276897 (U41809) - cyclin J [Drosophila melanogaster]	Contains protein domain (PF00134) - cyclin Cyclin	284683, 264689, 35698423, 284639
514	79813805 (1027, 1028)	Novel Protein sim. GBank gij1184790 (U46068) - von Ebner minor salivary gland protein [Mus musculus]	UNCLASSIFIED	29331830, 284909
515	79462591 (1029, 1030)			22278999, 284690
516	9862020 (1031, 1032)	Novel Protein sim. GBank gij1217400[pil]S65770 - maltotriose trehalose trehalohydrolase - Arthrobacter sp. (strain Q36)	amylase	284910

517	95292994 (1033, 1034)	Novel Protein sim. GBank gij2983605 (AE000725) - ribose 5 phosphate isomerase B [Aquifex aeolicus]		isomerase	265018, 264605, 264764, 264766, 264887, 264891, 264585
518	8491831 (1035, 1036)	Novel Protein sim. GBank gij854065[emb]CAA58337] - (X03413) U88 [Human herpesvirus 6]		UNCLASSIFIED	264487
519	91677888 (1037, 1038)	Novel Protein sim. GBank gij5689363[dbj]BAA83073.1] - (A0204075) B120 [Homo sapiens]	Contains protein domain (PF01388) - ARID DNA binding domain	dna_ma_bind	52644507, 22278997, 22278998, 60432049, 264259, 52845080, 29331824, 66714117, 60424269, 29331826, 35896052, 264905, 29331830, 66712502, 264511, 265007, 264591, 60432229, 33657402, 60433438, 21908754, 33109954, 52844298, 87188474, 87188559, 265017, 265018, 264604, 265019, 264681, 264448, 264389, 264288, 264885, 21908765, 21908766, 21906767, 21906769, 265021, 60170615, 33657023, 264692, 52645129, 33657109, 27486262, 27486264, 35893783, 18108370, 264629, 52844332, 56182323, 264639, 83373044, 18108385, 58528488, 60432113
520	79859188 (1039, 1040)	Novel Protein sim. GBank		transport	264789
521	11078821 (1041, 1042)	gij1169128[sp]P46839[CTPA_MYCLE - CATION-TRANSPORTING P-TYPE ATPASE A			264605
522	80435060 (1043, 1044)	Novel Protein sim. GBank	Contains protein domain (PF00294) - p18 family carbohydrate kinase	kinase	264805, 264768
523	18356013 (1045, 1046)	gij1172869[sp]P44331[RSK_HAEIN - RIBOKINASE Novel Protein sim. GBank gij12132243[pir]S81028 - hypothetical protein YPL236c - yeast (Saccharomyces cerevisiae)		UNCLASSIFIED	264629
524	80261805 (1047, 1048)	Novel Protein sim. GBank gij4033608[dbj]BAA35136] - (AB012308) B2HC [Anthracidaris crassispina]		ATPase associated	264082, 264596, 265011
525	78810046 (1049, 1050)	Novel Protein sim. GBank gij106610[emb]CAA21365] - (AL031866) ORF42, len=366 aa, similarity to an aminotransferase, in P95957 Sulfolobus solfataricus, (401 aa), 33.1% identity in 393 aa overlap, Fasta scores: opt:489, E(): 8.5e-24, in Q64602 R. norvegicus, (425 aa), 28.6% ident...			264807
526	38827630 (1051, 1052)	Novel Protein sim. GBank gij106610[emb]CAA21365] - (AL031866) ORF42, len=366 aa, similarity to an aminotransferase, in P95957 Sulfolobus solfataricus, (401 aa), 33.1% identity in 393 aa overlap, Fasta scores: opt:489, E(): 8.5e-24, in Q64602 R. norvegicus, (425 aa), 28.6% ident...		UNCLASSIFIED	264758
527	80504729 (1053, 1054)	Novel Protein sim. GBank		UNCLASSIFIED	264769
528	65484134 (1055, 1056)	gij731088[sp]P24215[UUA_ECOLI - MANNONATE DEHYDRATASE (D-MANNONATE HYDROLASE)]		UNCLASSIFIED	58182575, 265017, 265018
529	17936810 (1057, 1058)	Novel Protein sim. GBank gij42144[emb]CAA25200] - (X00513) NusA protein (nusA) [Escherichia coli]		hydrolase	265019
530	10887336 (1059, 1060)	Novel Protein sim. GBank gij42144[emb]CAA25200] - (X00513) NusA protein (nusA) [Escherichia coli]		UNCLASSIFIED	264687
531	80226576 (1061, 1062)	Novel Protein sim. GBank gij5262840[emb]CAB45758.1] - (AL080170) hypothetical protein [Homo sapiens]	Contains protein domain (PF00622) - SPRY domain	UNCLASSIFIED	264555, 264558, 264557, 264558, 18108385
532	90933444 (1063, 1064)	Novel Protein sim. GBank gij5262840[emb]CAB45758.1] - (AL080170) hypothetical protein [Homo sapiens]		UNCLASSIFIED	264488, 264490, 264259, 264592, 264760, 265021, 264690, 263976, 264558

533	87781531 (1085, 1068)	Novel Protein sim. GBank gi 4883636 gb AAD31593.1 AF112299 Integral inner nuclear membrane protein MAN1 [Homo sapiens]			284907, 284909, 284788, 35685917, 284630, 284555
534	82388284 (1087, 1068)	Novel Protein sim. GBank gi 2895352 emb CAA04608.1 (AJ001206) pep1 [Streptomyces coelicolor]	UNCLASSIFIED		284905, 285011, 284601, 284602, 284605, 284782, 284788, 285020, 284893, 284838, 284906
535	78641850 (1089, 1070)	Novel Protein sim. GBank gi 3878636 emb CAA88953 (Z49128) similar to cAMP-dependant protein kinase; cDNA EST EMBL: T00719 comes from this gene; cDNA EST yk485d8.3 comes from this gene; cDNA EST yk485d8.5 comes from this gene; cDNA EST yk49214.3 comes from this gene; cDNA EST y...	Contains protein domain (PF00089) - ATPase-associated Eukaryotic protein kinase domain		
536	78907207 (1071, 1072)	Novel Protein sim. GBank gi 2495628 sp P55757 YOH1_SERMA - HYPOTHETICAL 10.1 KD PROTEIN IN BIOA 5 REGION	reductase		18108376, 284805, 284906, 284807, 284909
537	84147448 (1073, 1074)	Novel Protein sim. GBank gi 134920 sp P21997 SSGP_VOLCA - SULFATED SURFACE GLYCOPROTEIN 185 (SSG 185)	collagen		285008, 284805, 85274791 28331822, 28331824, 28331825, 28331826, 28331827, 284908, 52644045, 33657402, 285017, 284782, 284883, 284288, 284885, 21908765, 35895783, 284558, 80170394, 284559, 22278002 284602, 285019
538	87821863 (1075, 1076)	Novel Protein sim. GBank gi 2498433 sp Q12341 HAT1_YEAST - HISTONE ACETYLTRANSFERASE	histone		
539	28398289 (1077, 1078)	Novel Protein sim. GBank gi 3882221 db BAA34480.1 (AB018303) KIAA0760 protein [Homo sapiens]	transcript factor		284893 18108394, 22278997, 22278998, 284259, 284112, 265009, 33657402, 55812038, 52848317, 285017, 21908765, 284893, 53811576, 284835, 58529486, 284566 284910, 265018, 284889, 284638, 284488
540	78637077 (1078, 1080)	Novel Protein sim. GBank gi 5042272 emb CAB44526.1 (AL078618) nuoF, NADH dehydrogenase subunit [Streptomyces coelicolor]	dehydrogenase		
541	87762268 (1081, 1082)	Novel Protein sim. GBank gi 2791398 emb CAA15994 (AL021184) hypothetical protein Rv1464 [Mycobacterium tuberculosis]	UNCLASSIFIED		284802, 284908
542	95205838 (1083, 1084)	Novel Protein sim. GBank gi 488721 gb AAD32237.1 AF14744 - (AF147449) penicillin binding protein 1B [Pseudomonas aeruginosa]	UNCLASSIFIED		284605
543	79798290 (1085, 1086)	Novel Protein sim. GBank gi 1350855 sp P19178 RPOC_PSEPU - DNA-DIRECTED RNA POLYMERASE BETA' CHAIN (TRANSCRIPTASE BETA' CHAIN) (RNA POLYMERASE BETA' SUBUNIT)	mapolymerase		284768, 284634, 284907, 284592, 284809 284600, 284602, 21908765
544	20437191 (1087, 1088)	Novel Protein sim. GBank gi 2330021 af AF019250 - kinesin- related protein; KRP; Costal2 [Drosophila melanogaster]			284604
545	80434504 (1089, 1090)				
546	80249018 (1091, 1092)				
547	11077563 (1093, 1094)				
548	82114938 (1095, 1096)		UNCLASSIFIED		284488, 284905, 284910, 284760, 284683, 284839, 284583, 284564

548	95421904 (1097, 1098)	Novel Protein sim. GBank gl 4337480 gb AAD18133 - (AF058195) neuroblastoma-amplified protein [Homo sapiens]		UNCLASSIFIED	264488, 65274572, 18108398, 22278995, 22278996, 22278997, 22278998, 22278999, 264259, 29331824, 66714117, 29331825, 29331826, 35896052, 265007, 265008, 264810, 264592, 33657402, 33109954, 265017, 265018, 265019, 18108351, 264448, 264764, 264389, 264288, 264766, 264686, 264688, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 264691, 33857023, 264692, 264693, 65274620, 52845129, 33657109, 27486261, 27486262, 27486264, 33657349, 55811576, 18108387, 60432113, 22279002, 264688
550	10886818 (1099, 1100)	Novel Protein sim. GBank		UNCLASSIFIED	264908, 264909, 264768
551	80439980 (1101, 1102)	gl 3122883 sp P84985 SYFB MYCTU - PHENYLALANYL-TRNA SYNTHETASE BETA CHAIN (PHENYLALANINE--TRNA LIGASE BETA CHAIN) [PHERS]		UNCLASSIFIED	264689, 264639, 264563, 55811957, 264628
552	94872870 (1103, 1104)	Novel Protein sim. GBank		UNCLASSIFIED	264906
553	80108002 (1105, 1106)	Novel Protein sim. GBank gl 5019771 gb AAD37857.1 AF13328 - (AF133263) histidine protein kinase-response regulator hybrid protein CvgSY [Pseudomonas syringae pv. syringae]	Contains protein domain (PF00008) - EGF-like domain	glycoprotein	264762
554	79618379 (1107, 1108)	Novel Protein sim. GBank		kinase	
555	78996347 (1109, 1110)	Novel Protein sim. GBank gl 131515 sp P02903 PTGA, SALT - PTS SYSTEM, GLUCOSE-SPECIFIC IIA COMPONENT (EIIA-GLC) (GLUCOSE-PERMEASE IIA COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, A COMPONENT) (EIIA-GLC)	Contains protein domain (PF00358) - phosphoenolpyruvate-dependent sugar phosphotransferase system, EIIA 1	transport	
556	20457127 (1111, 1112)	Novel Protein sim. GBank gl 3914014 sp P68380 IMFD, MYCTU - TRANSCRIPTION-REPAIR COUPLING FACTOR (TRCF)		transcript factor	264508, 264605, 264559
557	19523405 (1113, 1114)	Novel Protein sim. GBank gl 5042273 emb CAB44527.1 - (AL078618) nuoE, NADH dehydrogenase subunit [Siroptomyces coelicolor]		dehydrogenase	264488
558	20724429 (1115, 1116)	Novel Protein sim. GBank gl 1170933 sp P45331 METE, HAEIN - 5-METHYL-TETRAHYDROPTEROYL-TRIGLUTAMATE--HOMOCYSTEINE METHYLTRANSFERASE (METHIONINE SYNTHASE, VITAMIN-B12 INDEPENDENT ISOZYME) (COBALAMIN-INDEPENDENT METHIONINE SYNTHASE)		UNCLASSIFIED	264602
559	80084353 (1117, 1118)	Novel Protein sim. GBank gl 4980587 gb AAD35173.1 AEC00169 - (AEC001694) iron(III) ABC transporter, permease protein [Thermotoga maritima]		UNCLASSIFIED	264634

560	80066533 (1119, 1120)	Novel Protein sim. GBank gl 2492595 sp Q53183 Y4TR_RHISN - PROBABLE PEPTIDE ABC TRANSPORTER ATP-BINDING PROTEIN Y4TR	Contains protein domain (PF000005) - ABC transporter	transport	18108386, 284808, 284602, 284604, 18108374
561	20283187 (1121, 1122)			UNCLASSIFIED	284600
562	11698161 (1123, 1124)			UNCLASSIFIED	284889
563	79761420 (1125, 1126)	Novel Protein sim. GBank gl 4104925 (AF042276) - poly(hydroxycarboxylate) granule associated protein GA2 [Pseudomonas putida]		UNCLASSIFIED	284910, 284691
564	56716390 (1127, 1128)	Novel Protein sim. GBank gl 2792310 (AF040570) - unknown [Amycolatopsis mediterranei]		dehydrogenase	284592
565	58465818 (1128, 1130)	Novel Protein sim. GBank gl 3449294 db BAA32462 - (AB011532) MEGF8 [Rattus norvegicus]	Contains protein domain (PF000008) - EGF-like domain	synthase	265010
566	94323888 (1131, 1132)	Novel Protein sim. GBank gl 4539568 emb CAB38487.1 - (AL035636) putative helicase [Streptomyces coelicolor]		helicase	264909, 284510, 285008, 284910, 284758, 264600, 284602, 284604, 284605, 284789, 284687, 284689, 35895917, 284693, 65274620, 284486
567	79509955 (1133, 1134)	Novel Protein sim. GBank gl 100506 pr S17455 - Malate dehydrogenase (oxaloacetate-decarboxylating) (NADP+) (EC 1.1.1.40) - Flavaria trimeria (fragment)	Contains protein domain (PF00390) - Malic enzyme	UNCLASSIFIED	284681, 284691, 284593
568	94681793 (1135, 1136)	Novel Protein sim. GBank gl 3915643 sp O31212 RS2_STRCO - 30S RIBOSOMAL PROTEIN S2	Contains protein domain (PF00318) - Ribosomal protein S2	dehydrogenase	284689
569	39506897 (1137, 1138)	Novel Protein sim. GBank gl 115122 sp P21627 BRAD_PSEAE - HIGH-AFFINITY BRANCHED-CHAIN AMINO ACID TRANSPORT PROTEIN BRAD		ribosomal prot	284565
570	78375927 (1138, 1140)	Novel Protein sim. GBank		UNCLASSIFIED	18108378, 18108387, 284565
571	79793961 (1141, 1142)			transport	284907, 284909
572	38998838 (1143, 1144)	Novel Protein sim. GBank gl 4539223 emb CAB39881.1 - (AL049497) putative Integral membrane protein [Streptomyces coelicolor]		UNCLASSIFIED	284782
573	20715521 (1145, 1146)			UNCLASSIFIED	265007, 284601
574	13521592 (1147, 1148)	Novel Protein sim. GBank gl 118794 sp P10443 DP3A_ECOLI - DNA POLYMERASE III, ALPHA CHAIN			284638
575	13076418 (1149, 1150)			polymerase	284687
576	20482246 (1151, 1152)	Novel Protein sim. GBank gl 5457625 emb CAB49116.1 - (AJ248283) PAB2227 [Pyrococcus abyssi]			284605
577	68727.02 (1153, 1154)	Novel Protein sim. GBank gl 5042274 emb CAB44528.1 - (AL078618) nuoD, NADH dehydrogenase subunit [Streptomyces coelicolor]	Contains protein domain (PF00346) - Respiratory-chain NADH dehydrogenase, 49 Kd subunit	dehydrogenase	35696032, 284636
578	11804477 (1155, 1156)	Novel Protein sim. GBank			284638
579	11794723 (1157, 1158)	Novel Protein sim. GBank gl 1723081 sp Q11046 Y089_MYCTU - HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN CY50.09		transport	284682, 284556

580	80059417 (1158, 1160)					22278999, 35696052, 284555, 284558, 284558
581	79230833 (1161, 1162)				UNCLASSIFIED	284558
582	80049617 (1163, 1164)	Novel Protein sim. GBank gij3243131 (AF045777) - tilin [Drosophila melanogaster]	Contains protein domain (PF000047) - Immunoglobulin domain		struct	285008, 284564
583	79321392 (1165, 1166)	Novel Protein sim. GBank gij2501162spjP7728YAJR_ECOLI - HYPOTHETICAL 49.0 KD PROTEIN IN ABPA-CYOE INTERGENIC REGION			transport	285021, 284555, 284557
584	79845024 (1167, 1168)	Novel Protein sim. GBank gij3882221dbjBAA34470.11 - (AB018293) KIAA0750 protein [Homo sapiens]				284594
585	79581454 (1169, 1170)	Novel Protein sim. GBank gij3882221dbjBAA34470.11 - (AB018293) KIAA0750 protein [Homo sapiens]			UNCLASSIFIED	284488, 284908, 284768, 284687, 35688423
586	38277486 (1171, 1172)	Novel Protein sim. GBank gij4467250jemb CAB375751 - (AL035569) probable Glu-tRNA Gln amidotransferase subunit [Streptomyces coelicolor]			UNCLASSIFIED	285018, 284684, 21908769
587	80497359 (1173, 1174)	Novel Protein sim. GBank gij5689518dbjBAA83043.11 - (AB028014) KIAA1091 protein [Homo sapiens]			UNCLASSIFIED	284908, 265007
588	79557239 (1175, 1176)	Novel Protein sim. GBank gij5689518dbjBAA83043.11 - (AB028014) KIAA1091 protein [Homo sapiens]			hydrolase	284600, 284602, 284605, 284769, 284690, 284557
589	79805828 (1177, 1178)				UNCLASSIFIED	265020, 284692
590	79815628 (1179, 1180)	Novel Protein sim. GBank gij2143283jemb CAB093901 - (Z95972) rpoB [Mycobacterium tuberculosis]			UNCLASSIFIED	22278998, 284907, 284909, 284510, 285009, 285010, 284687, 284769, 35695917, 18108378, 284634, 284638, 284638
591	10313540 (1181, 1182)	Novel Protein sim. GBank gij2143283jemb CAB093901 - (Z95972) rpoB [Mycobacterium tuberculosis]			UNCLASSIFIED	284908, 284909
592	13869767 (1183, 1184)	Novel Protein sim. GBank gij4511983jgb AAD21543.11 - (AF088898) electroltransfer ubiqlnone oxidoreductase [Zymomonas mobilis]			mapolymerase	284691
593	82348698 (1185, 1186)	Novel Protein sim. GBank gij1272368 (U51896) - LjGE [Vibrio parahaemolyticus]			MHC	283972
594	20212392 (1187, 1188)	Novel Protein sim. GBank gij1272368 (U51896) - LjGE [Vibrio parahaemolyticus]			dehydrogenase	284511, 284762, 284769, 284488
595	10064084 (1189, 1190)	Novel Protein sim. GBank gij131490spjP20966jPTFB_ECOLI - PTS SYSTEM, FRUCTOSE-SPECIFIC IIBC COMPONENT (EIBC-FRU) (FRUCTOSE-PERMEASE IIBC COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, BC COMPONENT) (EII-FRU)			UNCLASSIFIED	284605
596	13085170 (1191, 1192)	Novel Protein sim. GBank gij125328spjP04951KDSB_ECOLI - 3-DEOXY-MANNO-OCTULONATE CYTIDYL TRANSFERASE (CMP-KDO SYNTHETASE) (CMP-2-KETO-3-DEOXYOCTULONIC ACID SYNTHETASE) (CKS)				284769
597	80259003 (1193, 1194)				UNCLASSIFIED	284838
598	94140216 (1195, 1196)				UNCLASSIFIED	284592
599	20385137 (1197, 1198)	Novel Protein sim. GBank gij125328spjP04951KDSB_ECOLI - 3-DEOXY-MANNO-OCTULONATE CYTIDYL TRANSFERASE (CMP-KDO SYNTHETASE) (CMP-2-KETO-3-DEOXYOCTULONIC ACID SYNTHETASE) (CKS)			UNCLASSIFIED	284758, 55810784, 284555, 284558, 284837, 83373044
600	10357683 (1199, 1200)	Novel Protein sim. GBank gij2127414jprj S60084 - hypohelical protein 2 - Corynebacterium glutamicum			UNCLASSIFIED	284803
601	79610404 (1201, 1202)	Novel Protein sim. GBank gij2127414jprj S60084 - hypohelical protein 2 - Corynebacterium glutamicum			UNCLASSIFIED	284908
					UNCLASSIFIED	284510

602	79250602 (1203, 1204)	Novel Protein sim. GBank gl 3522981 gb AAC34243.1 - (AC004411) putative pto kinase [Arabidopsis thaliana]	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	kinase	265007
603	11465067 (1205, 1206)			UNCLASSIFIED	264595
604	81875420 (1207, 1208)				264758
605	20436657 (1208, 1210)	Novel Protein sim. GBank gl 1175322 sp P44917 Y883_HAEIN - HYPOTHETICAL PROTEIN H10883		UNCLASSIFIED	264605
606	60334562 (1211, 1212)	Novel Protein sim. GBank gl 5020264 gb AAD38043.1 AF15136 - (AF151363) Cdc42 GTPase-activating protein [Mus musculus]		UNCLASSIFIED	264764
607	95381506 (1213, 1214)	Novel Protein sim. GBank gl 188864 (M74027) - mucin [Homo sapiens]		UNCLASSIFIED	264508, 264908, 55659542, 264682, 264687, 264689, 264534, 18108376, 35696423, 264638, 264555, 264638
608	11810888 (1215, 1216)			UNCLASSIFIED	264682
609	80064775 (1217, 1218)	Novel Protein sim. GBank gl 2498701 sp P5552 Y4LL_RHISN - HYPOTHETICAL 91.8 KD PROTEIN Y4LL	Contains protein domain (PF00989) - PAS domain	UNCLASSIFIED	264605
610	78829413 (1219, 1220)				264692
611	87588205 (1221, 1222)				264508, 264905, 264907, 264908, 264909, 264511, 264910, 264758, 264604, 264684, 264768, 264689, 264692, 264628, 264635, 264638, 264637, 264558
612	95287851 (1223, 1224)	Novel Protein sim. GBank gl 1877366 emb CAB07118 - (Z92772) recD [Mycobacterium tuberculosis]	Contains protein domain (PF01443) - Viral (Superfamily 1) RNA helicase	nuclease	264600, 264801, 264804, 264769, 264558, 264565
613	7523475 (1225, 1226)			UNCLASSIFIED	264389
614	79969348 (1227, 1228)	Novel Protein sim. GBank gl 511423 gb AAD40238.1 AF13670 - (AF136709) histidine Kinase YycG [Staphylococcus aureus]		kinase	18108372, 264563
615	39586998 (1229, 1230)	Novel Protein sim. GBank gl 1339950 db GAA12741 - (D85230) large subunit of NADH-dependent glutamate synthase [Plectonema boryanum]		synthase	264600, 264602, 264629
616	20465331 (1231, 1232)	Novel Protein sim. GBank gl 544367 sp P35673 GALE_ERWAM - UDP-GLUCOSE 4-EPIMERASE (GALACTOWALDENASE) (UDP-GALACTOSE 4-EPIMERASE)		isomerase	264605
617	91227222 (1233, 1234)	Novel Protein sim. GBank gl 2498087 sp Q60769 TNF3_MOUSE - TUMOR NECROSIS FACTOR, ALPHA-INDUCED PROTEIN 3 (PUTATIVE DNA BINDING PROTEIN A20) (ZINC FINGER PROTEIN A20)	Contains protein domain (PF00841) - Zn-finger in Ran binding protein and others.	Int	52845158, 21908765, 35696423, 21908768, 21908769, 22278994, 35696286, 22278998, 265020, 265021, 265007, 265008, 264638, 52844150, 33657023, 264692, 264693, 29331822, 29331824, 55812036, 83373044, 58182181, 60424269, 68714117, 29331825, 33657109, 29331826, 33657182, 29331827, 35696052, 29331828, 27486282, 33657349, 56528488, 285018, 265019, 22279002, 704482, 264448, 29331830, 68712502, 284909

618	20632843 (1235, 1236)	Novel Protein sim. GBank gij5459388[emb]CAB50746.1] - (AL098639) putative aminotransferase [Streptomyces coelicolor]		isomerase	264603
619	91227224 (1237, 1238)				56994075, 28331826, 33656970, 265008, 33857402, 33109954, 87168559, 284448, 18108374, 83373044
620	81183143 (1239, 1240)	Novel Protein sim. GBank gij464333[sp]Q05922[us2_MOUSE - DUAL SPECIFICITY PROTEIN PHOSPHATASE 2 (DUAL SPECIFICITY PROTEIN PHOSPHATASE PAC-1)]	phosphatase		28148498, 264758, 284369, 29148627
621	80239251 (1241, 1242)	Novel Protein sim. GBank gij2633557[emb]CAB1306[0] - (Z69110) yJdF [Bacillus subtilis]	UNCLASSIFIED		264556, 284558, 264639
622	20456427 (1243, 1244)	Novel Protein sim. GBank gij1857710[gb]A4848482] - (U87224) contactin associated protein [Rattus norvegicus]	UNCLASSIFIED		264605
623	10131798 (1245, 1246)	Novel Protein sim. GBank gij1705703[sp]P52225[CCMF_PSEFL - CYTOCHROME C-TYPE BIOGENESIS PROTEIN CYCK	Contains protein domain (PF00054) - Laminin G domain	flavinin	284906
624	19534127 (1247, 1248)	Novel Protein sim. GBank gij1705703[sp]P52225[CCMF_PSEFL - CYTOCHROME C-TYPE BIOGENESIS PROTEIN CYCK	cytochrome		264596
625	13084619 (1249, 1250)	Novel Protein sim. GBank gij2894252[emb]CAA17114.1] - (AL021841) hypothetical protein Rv3342 [Mycobacterium tuberculosis]	UNCLASSIFIED		264688
626	80662603 (1251, 1252)	Novel Protein sim. GBank gij416592[sp]P32323[JAGA1_YEAST - A-AGGLUTININ ATTACHMENT SUBUNIT PRECURSOR	UNCLASSIFIED		29331822, 264905, 284908, 33657023, 33657109, 264558
627	80255457 (1253, 1254)	Novel Protein sim. GBank gij3098418 (AF040944) - P140 [Mus musculus]	UNCLASSIFIED		18108394, 264112, 264593, 265022, 264635
628	80077096 (1255, 1256)	Novel Protein sim. GBank gij1711543[sp]P50526[SSP1_SCHPO - SERINE/THREONINE-PROTEIN KINASE SSP1	Contains protein domain (PF00089) - Eukaryotic protein kinase domain	kinase	264600
629	79851802 (1257, 1258)	Novel Protein sim. GBank gij1143204 (U34305) - ORF2; Method: conceptual translation supplied by author. [Shigella sonnei]	isomerase		284908, 264907
630	39565156 (1259, 1260)	Novel Protein sim. GBank gij3236368 (AF064748) - S3-12 [Mus musculus]	UNCLASSIFIED		264480
631	20598718 (1261, 1262)	Novel Protein sim. GBank gij140687[sp]P11668[YGGB_ECOLI - HYPOTHETICAL 30.9 KD PROTEIN IN SBM-FBA INTERGENIC REGION (ORF 4) (F286)]			263978
632	27843890 (1263, 1264)		UNCLASSIFIED		264908, 264600, 264605, 264769, 264689, 264488
633	80477772 (1265, 1266)		UNCLASSIFIED		264769
634	17938808 (1267, 1268)		UNCLASSIFIED		265019
635	79574508 (1269, 1270)		UNCLASSIFIED		264689
636	79910981 (1271, 1272)		UNCLASSIFIED		264596, 264782, 264693

637	82455798 (1273, 1274)	Novel Protein sim. GBank gl 2326739 emb CAB10953 - (Z98288) recN [Mycobacterium tuberculosis]	nuclease	284906, 284907, 284510, 284511, 284601, 284602, 284603, 284604, 284605, 18108351, 284782, 284786, 284687, 284789, 284689, 35695917, 284693, 284634, 284638, 284639, 284559, 18108385, 284638
638	14997457 (1275, 1276)	Novel Protein sim. GBank gl 4678662 emb CAB1074.1 - (AL049645) putative large ATP-binding protein [Streptomyces coelicolor]		
639	80204210 (1277, 1278)	Novel Protein sim. GBank gl 4589628 dbj BAA76836.1 - (AB023209) KIAA0992 protein [Homo sapiens]	struct	284112, 263974
640	17929579 (1278, 1280)	Novel Protein sim. GBank gl 1432083 (U60981) - homolog to Skp1p, an evolutionarily conserved kinetochore protein in budding yeast [Arabidopsis thaliana]	Contains protein domain (PF01466) - Skp1 family	285009, 265010
641	78638398 (1281, 1282)		UNCLASSIFIED	284693
642	19898737 (1283, 1284)		UNCLASSIFIED	284565
643	61516220 (1285, 1286)		UNCLASSIFIED	284906, 284908, 284758, 284288, 284632, 284635, 284639, 284584
644	11751367 (1287, 1288)		UNCLASSIFIED	284684
645	95010907 (1289, 1290)		UNCLASSIFIED	284906, 284762, 284693, 284639, 284559
646	80069083 (1291, 1292)		UNCLASSIFIED	284595, 284568
647	80257085 (1293, 1294)		transcript factor	284908, 284591
648	80077428 (1295, 1296)	Novel Protein sim. GBank gl 4507813 ref NP_003738.1 TNKS - TANKYRASE	UNCLASSIFIED	284600
649	80247447 (1297, 1298)	Novel Protein sim. GBank gl 1044863 bbs 169846 - prolamine [Monodelphis domestica, gonads, Peptide, 108 aa]	UNCLASSIFIED	263978
650	11788316 (1299, 1300)		UNCLASSIFIED	284688
651	11776932 (1301, 1302)	Novel Protein sim. GBank gl 1346916 sp P12283 PURA_ECOLI - ADENYLOSUCCINATE SYNTHETASE (IMP-ASPARTATE LIGASE)		284602, 284638
652	85518704 (1303, 1304)		UNCLASSIFIED	284905, 284907, 284809, 283978, 284637
653	82124947 (1305, 1306)	Novel Protein sim. GBank gl 1722977 sp Q10638 Y03C_MYCTU - HYPOTHETICAL 82.8 KD PROTEIN CY130.12C	UNCLASSIFIED	22278998, 284510, 284511, 284512, 284593, 21906754, 284603, 284760, 18108376, 284556
654	95010589 (1307, 1308)		UNCLASSIFIED	284908, 284595, 284632
655	78320692 (1309, 1310)	Novel Protein sim. GBank gl 130327 sp P26847 PLSC_ECOLI - 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE (1-AGP ACYLTRANSFERASE) (1-AGPAT) (LYSOPHOSPHATIDIC ACID ACYLTRANSFERASE) (LPAAT)	transferrase	284592
656	80416739 (1311, 1312)		Contains protein domain (PF01553) - Acyltransferase	
657	20811010 (1313, 1314)		UNCLASSIFIED	284602, 284605, 284766, 284691
			UNCLASSIFIED	284557, 284558

658	87761815 (1315, 1316)	Novel Protein sim. GBank gij5689493dbj BAA83030.1 - (AB029001) KIAA1078 protein (Homo sapiens)	UNCLASSIFIED	22278996, 60432049, 29331822, 29331824, 29331828, 265007, 265009, 33657402, 33657084, 265017, 264448, 21906765, 21906766, 263987, 20281149, 18108370, 18108374, 264482
659	87718663 (1317, 1318)	Novel Protein sim. GBank gij2137872jpr j48724 - zinc finger protein PZF - mouse	transcriptfactor	22278999, 60432049, 66714117, 29331827, 265007, 264766, 56181562, 18108359, 18108365, 18108370, 18108381
660	81897922 (1319, 1320)	Novel Protein sim. GBank	UNCLASSIFIED	284757
661	80026023 (1321, 1322)	Novel Protein sim. GBank gij134180 spP15401 SACY_BACSU - LEVANSUCRASE AND SUCRASE SYNTHESIS OPERON ANTITERMINATOR	UNCLASSIFIED	264510, 265009, 264600, 264602, 264603, 264604, 264605, 32833986, 18108378, 264636, 18108387, 22279000
662	20463731 (1323, 1324)	Novel Protein sim. GBank gij4545229jgb AAD22450.1 AF11618 - (AF116183) SecA homolog (Actinobacillus actinomycetemcomitans)	UNCLASSIFIED	264605
663	20626080 (1325, 1326)	Novel Protein sim. GBank gij5689250jdbj BAA82881.1 - (AB024335) similar to orf5 (Comamonas testosteroni)	dehydrogenase	284605
664	80508512 (1327, 1328)	Novel Protein sim. GBank gij1652848jdbj BAA17768 - (D90909) DNA photolyase (Synecococcus sp.)	UNCLASSIFIED	264769
665	80079053 (1329, 1330)	Novel Protein sim. GBank gij116841 spP21640 COBJ_PSEDE - PRECORRIN-3 METHYLTRANSFERASE (PRECORRIN-3 METHYLTRANSFERASE) (PRECORRIN-3 METHYLTRANSFERASE) (PRECORRIN-3 METHYLTRANSFERASE)	isomerase	264600
666	79603142 (1331, 1332)	Novel Protein sim. GBank gij3261829jemb CAB10927 - (Z98260) hypothetical protein Rv1230c (Mycobacterium tuberculosis)	glycoprotein	284907, 265007
667	84631802 (1333, 1334)	Novel Protein sim. GBank gij5688851jdbj BAA82702.1 - (AB017438) Orf5 (Streptomyces coelicolor)	UNCLASSIFIED	284689, 264602, 264593
668	82051891 (1335, 1336)	Novel Protein sim. GBank gij3581853jemb CAA20809 - (AL031541) 50S ribosomal protein L20 (Streptomyces coelicolor)	ribosomalprot	264905, 264906, 264908, 264600, 264601, 264603, 264605, 264760, 264689, 264636, 264638, 264639
669	12967154 (1337, 1338)	Novel Protein sim. GBank gij2582531 (AF026444) - 2- isopropylmalate synthase (Streptomyces coelicolor)	UNCLASSIFIED	264637
670	80238548 (1339, 1340)	Novel Protein sim. GBank gij2582531 (AF026444) - 2- isopropylmalate synthase (Streptomyces coelicolor)	synthase	264905, 264906, 264908, 264601, 264762, 264766, 264689, 264636, 18108385, 264486
671	78601388 (1341, 1342)	Novel Protein sim. GBank gij2114430 (U92703) - Olf-1/EBF-like-3 transcription factor (Mus musculus)	UNCLASSIFIED	264690, 264692, 264693, 264636, 18108387
672	78634371 (1343, 1344)	Novel Protein sim. GBank gij4589285jgb AAD28430.1 AF13515 - (AF135154) feric aicallign aldophore receptor (Bordetella pertussis)	transcriptfactor	264910, 265017
673	82285788 (1345, 1346)	Novel Protein sim. GBank gij4589285jgb AAD28430.1 AF13515 - (AF135154) feric aicallign aldophore receptor (Bordetella pertussis)		284759
674	79189259 (1347, 1348)	Novel Protein sim. GBank gij4589285jgb AAD28430.1 AF13515 - (AF135154) feric aicallign aldophore receptor (Bordetella pertussis)	UNCLASSIFIED	264629

675	87895970 (1349, 1350)	Novel Protein sim. GBank gi 4980755 gb AAD33347.1 AE0011708 D- alanine--D-alanine ligase [Thermotoga maritima]	Contains protein domain (PF01820) - D-alta D-alta ligase	UNCLASSIFIED	284488, 22278999, 68714117, 284508, 284511, 285008, 80433438, 284600, 284601, 284602, 284603, 284604, 284605, 284782, 284687, 284789, 80431802, 18108374, 284638, 284638 285010
676	78898607 (1351, 1352)	Novel Protein sim. GBank gi 1723566 sp Q10479 YDF7_SCHPO - PUTATIVE GLUCOSYLTRANSFERASE C17C9.07			
677	21644312 (1353, 1354)	Novel Protein sim. GBank gi 887208 (U03978) - dynein heavy chain isoform 5C [Tripleneustes grallia]		ATPase-associated	284591, 284632
678	84225200 (1355, 1356)	Novel Protein sim. GBank gi 1586274 pf 2203365A - laminin alpha5 [Mus musculus]	Contains protein domain (PF00053) - Laminin EGF-like (Domains III and V)	laminin	284758, 284682, 284557
679	78868855 (1357, 1358)	Novel Protein sim. GBank gi 3928723 emb CAA22219 - (AL034355) putative ABC transporter [Streptomyces coelicolor]		UNCLASSIFIED	22278998, 284693
680	20726424 (1359, 1360)	Novel Protein sim. GBank			
681	94322017 (1361, 1362)	Novel Protein sim. GBank gi 5174493 ref NP_006050.1 pLAMC - laminin, gamma 3	Contains protein domain (PF00053) - Laminin EGF-like (Domains III and V)	laminin	284600, 284602 284102, 284907, 284908, 285008, 284693, 283972, 83373044, 284568
682	11392478 (1363, 1364)	Novel Protein sim. GBank		UNCLASSIFIED	284595
683	80083680 (1365, 1366)	Novel Protein sim. GBank gi 4758208 ref NP_004081.1 pDUSP - dual specificity phosphatase 3 (vaccinia virus phosphatase VH1-related)	Contains protein domain (PF00782) - Dual specificity phosphatase, catalytic domain	phosphatase	284634
684	20465387 (1367, 1368)	Novel Protein sim. GBank gi 5420387 emb CAB46878.1 - (AJ213459) proteophosphoglycan [Leishmania major]			284605
685	80246735 (1369, 1370)				
686	78208608 (1371, 1372)				284809, 283987, 283981
687	80085829 (1373, 1374)				284631
688	78853412 (1375, 1376)	Novel Protein sim. GBank gi 2688962 (AF027768) - LspA [Serattia marcescens]		peptidase	284693, 284635 284907, 284638
689	88084258 (1377, 1378)	Novel Protein sim. GBank gi 3046931 (AF049330) - PPAR gamma coactivator [Mus musculus]	Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	UNCLASSIFIED	284908, 284907, 285007, 285009, 60433438, 21906754, 284760, 18108338, 21908788, 21908789, 285021, 18108381, 283974, 18108379, 284557, 18108385, 22278002
690	80389750 (1378, 1380)	Novel Protein sim. GBank gi 2498941 sp Q15428 SP62_HUMAN - SPICEOSOME ASSOCIATED PROTEIN 62 (SAP 62) (SF3A68)		UNCLASSIFIED	284510, 284511, 284784, 284769
691	81854392 (1381, 1382)			UNCLASSIFIED	284757
692	83608938 (1383, 1384)	Novel Protein sim. GBank gi 5420387 emb CAB46878.1 - (AJ243459) proteophosphoglycan [Leishmania major]	Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger)	UNCLASSIFIED	35812038, 55811957, 285018, 55811150, 18108351, 284808, 80431528, 284594
693	78566118 (1385, 1386)	Novel Protein sim. GBank gi 854065 emb CAA58337 - (X83413) U88 [Human herpesvirus 6]		UNCLASSIFIED	284635
694	82455983 (1387, 1388)	Novel Protein sim. GBank gi 267327 sp Q01033 VG48_HSVSA - HYPOTHETICAL GENE 48 PROTEIN			22278998, 284510, 284602, 284603, 284762, 284687, 284769, 284693

695	94147849 (1389, 1390)	Novel Protein sim. GBank gl 4468339 emb CAB38059.1 - (AJ010901) MUC4 [Homo sapiens]	Contains protein domain (PF00094) - von Willebrand factor type D domain	UNCLASSIFIED	56182575, 264509, 264905, 264907, 29331830, 264908, 264909, 264511, 285007, 264910, 264758, 264764, 264288, 65274791, 264905, 264595
696	78830882 (1391, 1392)	Novel Protein sim. GBank gl 2649950 (AE001058) - glutamine ABC transporter; ATP-binding protein (glnQ) [Archaeoglobus fulgidus]	Contains protein domain (PF00005) - ABC transporter	transport	264682
697	11767889 (1393, 1394)	Novel Protein sim. GBank gl 1731343 sp Q10894 YY25_MYCTU - HYPOTHETICAL 24.4 KD PROTEIN CY48.25	Contains protein domain (PF01836) - Transposase	UNCLASSIFIED	264682
698	66695862 (1395, 1396)			UNCLASSIFIED	264688, 35695917
699	78582558 (1397, 1398)			UNCLASSIFIED	264682
700	78639098 (1399, 1400)			UNCLASSIFIED	264693
701	80230242 (1401, 1402)	Novel Protein sim. GBank gl 1001236 dbj BAA10477 - (D84003) hypothetical protein [Synecocystis sp.]		UNCLASSIFIED	264488, 264510, 264511, 264602, 264605, 264689
702	78814789 (1403, 1404)	Novel Protein sim. GBank gl 2488935 sp Q48338 SOXG_CORSP - SARCOSINE OXIDASE GAMMA SUBUNIT		UNCLASSIFIED	264809
703	20446620 (1405, 1406)	Novel Protein sim. GBank gl 3150513 (AF067219) - contains similarity to the kelch/MIPP family [Caenorhabditis elegans]		oxidase	264604
704	94312224 (1407, 1408)	Novel Protein sim. GBank gl 421091 pir J530730 - hypothetical protein o208 - Escherichia coli	Contains protein domain (PF01344) - Kelch motif	UNCLASSIFIED	264288, 56181562, 33657109, 264629, 55811578
705	17932141 (1409, 1410)	Novel Protein sim. GBank gl 3024672 sp Q55790 Y074_SYNY3 - HYPOTHETICAL 52.8 KD PROTEIN SLR0074		UNCLASSIFIED	265006
706	20288062 (1411, 1412)	Novel Protein sim. GBank gl 3420608 gb AAC31807.1 - (AF075709) ABC transporter ATP-binding subunit [Pseudomonas putida]		transport	264600
707	20638065 (1413, 1414)	Novel Protein sim. GBank gl 3649741 emb CAA03985 - (AJ000281) mucin [Homo sapiens]		struct	264603
708	88001439 (1417, 1418)	Novel Protein sim. GBank gl 3080425 emb CAA18744.1 - (AL022604) putative protein [Arabidopsis thaliana]			264601, 264692
709	11356683 (1419, 1420)	Novel Protein sim. GBank gl 4758686 ref NP_002323.1 pLRP1 - low density lipoprotein related protein 1 (alpha-2-macroglobulin receptor)			18108398, 264637, 264808, 264809
710	17931418 (1421, 1422)	Novel Protein sim. GBank gl 1703286 sp Q11056 AMIZ_MYCTU - PUTATIVE AMIDASE CY50.19C	Contains protein domain (PF00058) - apolipoprotein		264369
711	80258164 (1423, 1424)	Novel Protein sim. GBank gl 4502351 ref NP_001692.1 pBAAT - bile acid Coenzyme A: amino acid N-acetyltransferase; glycine N-choloyltransferase	Low-density lipoprotein receptor repeat class B		265019
712	79263126 (1425, 1426)	Novel Protein sim. GBank gl 4502351 ref NP_001692.1 pBAAT - bile acid Coenzyme A: amino acid N-acetyltransferase; glycine N-choloyltransferase		hydrolase	264591
713	79263126 (1425, 1426)	Novel Protein sim. GBank gl 4502351 ref NP_001692.1 pBAAT - bile acid Coenzyme A: amino acid N-acetyltransferase; glycine N-choloyltransferase			264906, 264807
714	27847651 (1427, 1428)	Novel Protein sim. GBank gl 4502351 ref NP_001692.1 pBAAT - bile acid Coenzyme A: amino acid N-acetyltransferase; glycine N-choloyltransferase			264508, 264555

715	79639423 (1429, 1430)	Novel Protein sim. GBank gij1789035 (AE000332) - ori. hypothetical protein [Escherichia coli]		UNCLASSIFIED	284907	
716	79559072 (1431, 1432)				284692	
717	79491842 (1433, 1434)	Novel Protein sim. GBank gij249074spj55653[GABD_RHISN - PROBABLE SUCCINATE-SEMIALDEHYDE DEHYDROGENASE (NADP+)] [SSDH]		dehydrogenase	284636	
718	94319658 (1435, 1436)	Novel Protein sim. GBank gij3873679[embjCAA94886] - (Z71178) similar to pro-collagen domains; cDNA EST EMBL:D27978 comes from this gene; cDNA EST EMBL:D27977 comes from this gene; cDNA EST EMBL:D34199 comes from this gene; cDNA EST EMBL:D64392 comes from this gene; cDNA EST EMBL...	Contains protein domain (PF00089) - von Willebrand factor type C domain	kinase	18108392, 22278994, 22278998, 265008, 265018, 264681, 18108354, 264684, 264685, 264686, 264687, 264689, 21906789, 18108361, 264691, 264692, 55810764, 264635, 18108381, 18108382, 83373044, 18108388	
719	17879564 (1437, 1438)	Novel Protein sim. GBank gij2104302[embjCAB08631] - (Z95387) hypothetical protein Rv2611c [Mycobacterium tuberculosis]		UNCLASSIFIED	285011	
720	79841884 (1439, 1440)				264608	
721	15020180 (1441, 1442)	Novel Protein sim. GBank gij123530spjP04929[HRPX_PLALO - HISTIDINE-RICH GLYCOPROTEIN PRECURSOR		UNCLASSIFIED	264629	
722	9862603 (1443, 1444)	Novel Protein sim. GBank gij498253 (U02372) - integrase [Vibrio cholerae]			264910	
723	19755599 (1445, 1446)	Novel Protein sim. GBank gij2253054[embjCAB10705] - (Z97559) hypothetical protein Rv2114 [Mycobacterium tuberculosis]		UNCLASSIFIED	264691	
724	10126484 (1447, 1448)	Novel Protein sim. GBank gij4083015 (AF083061) - protease PrtA [Pseudomonas fluorescens]	Contains protein domain (PF00353) - Hemolysin-type calcium-binding proteins	protease	264909	
725	78878879 (1449, 1450)			UNCLASSIFIED	264905, 284907	
726	13086282 (1451, 1452)			UNCLASSIFIED	264636	
727	13522872 (1453, 1454)				264634	
728	20268471 (1455, 1456)	Novel Protein sim. GBank gij2633910[embjCAB13411] - (Z99112) similar to hypothetical proteins [Bacillus subtilis]			264567	
729	11293753 (1457, 1458)			UNCLASSIFIED	264490	
730	18900373 (1459, 1460)	Novel Protein sim. GBank gij2494680spjQ45291[GALE_BRELA - UDP-GLUCOSE 4- EPIMERASE (GALACTOWALDENASE) (UDP- GALACTOSE 4-EPIMERASE)		isomerase	264564	
731	60058750 (1461, 1462)	Novel Protein sim. GBank gij1146192 (L47838) - putative [Bacillus subtilis]		UNCLASSIFIED	264605	
732	80258175 (1463, 1464)	Novel Protein sim. GBank gij1168398spjP46881[AIIP2_YEAST - ACTIN INTERACTING PROTEIN 2		struct	264591, 264594, 264595	
733	20446839 (1465, 1466)			UNCLASSIFIED	264604	
734	20435987 (1467, 1468)	Novel Protein sim. GBank gij3184080[embjCAA19336] - (AL023781) hypothetical protein [Schizosaccharomyces pombe]		ubiquitin	264604	

735	11607959 (1469, 1470)	Novel Protein sim. GBank gi 401582 sp P27432 YICE_ECOLI - HYPOTHETICAL 48.9 KD PROTEIN IN GLTS-SELF INTERGENIC REGION			264594	
736	10879734 (1471, 1472)	Novel Protein sim. GBank gi 400831 sp P31135 POTH_ECOLI - PUTRESCINE TRANSPORT SYSTEM PERMEASE PROTEIN POTH	Contains protein domain (PF00528) - Binding-protein-dependent transport systems inner membrane component	transport	264636	
737	76945340 (1473, 1474)		Contains protein domain (PF00815) - Regulator of G protein signaling domain	UNCLASSIFIED	265020	
738	17895353 (1475, 1476)				265008	
739	79833670 (1477, 1478)	Novel Protein sim. GBank gi 2506867 sp P33225 TORA_ECOLI - TRIMETHYLAMINE- N-OXIDE REDUCTASE PRECURSOR (TMAO REDUCTASE) (TRIMETHYLAMINE OXIDASE)		oxidase	264910	
740	19881557 (1479, 1480)	Novel Protein sim. GBank gi 3261828 emb CAB10925 - (Z98260) mmp [Mycobacterium tuberculosis]	Contains protein domain (PF01863) - Domain of unknown function	UNCLASSIFIED	264907, 264764, 264634, 264637 264688, 35698286, 264510, 264908, 18108382	
741	79827273 (1481, 1482)	Novel Protein sim. GBank gi 3877494 emb CAA88472.1 - (Z48593) ATP binding protein with similarity to the CDC48/PAS1/SEC18 family; cDNA EST EMBL:D65037 comes from this gene; cDNA EST EMBL:D68340 comes from this gene; cDNA EST EMBL:D65048 comes from this gene; cDNA EST EMBL:D6845...		UNCLASSIFIED	29331822, 264910, 264762	
742	82393795 (1483, 1484)					
743	82300051 (1485, 1486)	Novel Protein sim. GBank gi 127420 sp P19888 MTBA_BACAR - MODIFICATION METHYLASE BANI (CYTOSINE-SPECIFIC METHYLTRANSFERASE BANI) (M.BANI)	Contains protein domain (PF00145) - C-5 cytosine-specific DNA methylase	UNCLASSIFIED	264488, 264259, 264508, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 264512, 265008, 265009, 264910, 264591, 264596, 264759, 265010, 265011, 18108351, 264783, 264288, 264768, 264768, 264693, 18108370, 264629, 18108372, 264630, 264631, 264634, 264558, 18108385, 264482, 264584, 264567 18108397, 264511, 264690, 264628, 264638, 264692, 264639, 264766 264906	
744	80230421 (1487, 1488)			glycoprotein		
745	9841963 (1489, 1490)	Novel Protein sim. GBank gi 76921 pir J504846 - UDP-N- acetylmutamoylalanine-D-glutamate-2, 6-diaminopimelate--O- alanine-D-alanine ligase (EC 6.3.2.15) precursor - Escherichia coli				
746	11073229 (1491, 1492)	Novel Protein sim. GBank gi 3366354 (AF074705) - pyochelin synthetase [Pseudomonas aeruginosa]		synthase	264600	
747	94322044 (1493, 1494)	Novel Protein sim. GBank gi 2887411 db BAA24848 - (AB007878) KIAA0419 [Homo sapiens]	Contains protein domain (PF00018) - SH3 domain	oxidase	56714117, 264905, 264509, 264906, 264907, 264908, 264909, 264511, 264910, 265011, 264681, 264288, 264766, 264687, 264768, 264769, 21906768, 35695917, 264691, 264693, 264628, 264634, 264635, 264639, 56182323, 83373044	
748	11617923 (1495, 1496)				264690	

749	20469119 (1497, 1498)	Novel Protein sim. GBank		UNCLASSIFIED	264604
750	20298427 (1499, 1500)	gi 1169727 sp P44948 FPG_HAEIN - FORMAMIDOPYRIMIDINE-DNA GLYCOSYLASE (FAPY-DNA GLYCOSYLASE)			264600
751	21636169 (1501, 1502)	Novel Protein sim. GBank gi 5380089 gb AAD42851.1 AF15968 - (AF159689) serine/threonine kinase PKN3 [Mycobacterium xenopus]	Contains protein domain (PF00089) - kinase Eukaryotic protein kinase domain		264605, 264559
752	82450386 (1503, 1504)	Novel Protein sim. GBank gi 1168682 sp P44428 BIOA_HAEIN - ADENOSYLMETHIONINE-8-AMINO-7-OXONONANOATE AMINOTRANSFERASE (7.8-DIAMINO-PELARGONIC ACID AMINOTRANSFERASE) (DAPA AMINOTRANSFERASE)	Contains protein domain (PF00202) - Aminotransferases class-III pyridoxal phosphate	UNCLASSIFIED	264508, 264907, 264510, 265011, 264782, 264689, 35695855, 264638, 18108387
753	80508718 (1505, 1506)	Novel Protein sim. GBank gi 2851530 sp P32399 YHGE_BACSU - HYPOTHETICAL 84.1 KD PROTEIN IN HEMY-GLTT INTERGENIC REGION (ORF8)		UNCLASSIFIED	264909, 264600, 264602, 264604, 264760, 264769, 264634
754	95083741 (1507, 1508)				
755	80185449 (1509, 1510)	Novel Protein sim. GBank gi 3449276 emb CAA20420 - (AL031317) putative dehydrogenase [Streptomyces coelicolor]		UNCLASSIFIED	264508, 264908, 264907, 264908, 264909, 264759, 264602, 264784, 264789, 264828, 264628, 264630, 264632, 264634, 264635, 264637, 264638, 83373044, 18108385
756	94631688 (1511, 1512)			UNCLASSIFIED	264448, 264690
757	79468533 (1513, 1514)				264769, 264689, 264638, 264639
758	78983178 (1515, 1516)	Novel Protein sim. GBank gi 4580331 emb CAB40107.1 - (AJ001206) putative glycogen debranching enzyme [Streptomyces coelicolor]		UNCLASSIFIED	264682, 264685
759	79475667 (1517, 1518)	Novel Protein sim. GBank gi 2911858 (AF047659) - No definition line found [Caenorhabditis elegans]	amylase		265007, 18108387, 265007, 18108387
760	87628888 (1519, 1520)	Novel Protein sim. GBank gi 3451312 emb CAA20449 - (AL031324) membrane alase [Schizosaccharomyces pombe]		UNCLASSIFIED	264684, 264686
761	79877988 (1521, 1522)		Contains protein domain (PF00122) - transport E1-E2 ATPase		28331822, 264908, 52844045, 56182435, 60170831, 21908754, 285017, 285019, 284681, 264687, 264688, 21906768, 21908768, 285020, 285021, 285022, 264635, 22279000
762	80023583 (1523, 1524)	Novel Protein sim. GBank gi 3327158 dbj BAA31847 - (AB014572) KIAA0872 protein [Homo sapiens]		UNCLASSIFIED	264768
763	20284813 (1525, 1526)	Novel Protein sim. GBank gi 4981268 gb AAD35822.1 AE00174 - (AE001744) lipopolysaccharide core biosynthesis protein KdIB [Thermotoga maritima]		UNCLASSIFIED	264907, 264593, 265020
764	39515024 (1527, 1528)				264600
					264603

765	80025347 (1529, 1530)	Novel Protein sim. GBank gij3845093 (AE001371) - erythrocyte membrane protein PIEMP3 [Plasmodium falciparum]		struct	284905, 284908, 284594, 284688, 33657023
766	82417404 (1531, 1532)				
767	10298742 (1533, 1534)	Novel Protein sim. GBank gij341112 [pir]S40827 - hypothetical protein o300 - Escherichia coli		UNCLASSIFIED	284605, 284762, 18108374
768	79416080 (1535, 1536)				284907
769	80086554 (1537, 1538)	Novel Protein sim. GBank gij2882501 [emb]CAA06184 - (AJ004832) neuropathy target esterase [Homo sapiens]		UNCLASSIFIED	284592, 284595
770	80417847 (1539, 1540)	Novel Protein sim. GBank gij283437 [pir]S27850 - hypothetical protein - Trypanosoma cruzi (fragment)		esterase	55810764, 264559
771	95329509 (1541, 1542)	Novel Protein sim. GBank gij4789004 [gb]AAD29715.1 [AF140598] ring-box protein 1 [Homo sapiens]	Contains protein domain (PF000097) - Zinc finger, C3HC4 type (RING finger)	UNCLASSIFIED	284905, 284907, 284828, 284909, 265010, 284766, 284828, 284629, 284634, 284638, 264555
772	78971362 (1543, 1544)			UNCLASSIFIED	56182575, 35586286, 264259, 66714117, 264107, 66712502, 56182435, 264112, 55812038, 87168559, 284288, 21906788, 33657023, 65274620, 65274791, 18108381
773	78945383 (1545, 1546)			UNCLASSIFIED	284910
774	78858129 (1547, 1548)	Novel Protein sim. GBank gij5531324 [emb]CAB51045.1 - (AJ009579) putative alkane 1-monooxygenase [Pseudomonas fluorescens]		UNCLASSIFIED	285020
775	20620141 (1549, 1550)			UNCLASSIFIED	284909
776	78942693 (1551, 1552)		Contains protein domain (PF01006) - Hepatitis C virus non-structural protein NS4a	UNCLASSIFIED	284555
777	79960378 (1553, 1554)	Novel Protein sim. GBank gij4505461 [ref]NP_003624.1 [p]NRPB - nuclear restricted protein, BTB domain-like (brain)	Contains protein domain (PF01344) - Kelch motif	protease	21908754, 285020, 60170615, 284691
778	20691310 (1555, 1556)			UNCLASSIFIED	264511
779	80054024 (1557, 1558)				284603
780	85288987 (1559, 1560)	Novel Protein sim. GBank gij1144520 (U34956) - phosphoribosylformylglycinamide synthase [Mycobacterium tuberculosis]		synthase	284907, 284600, 284601, 284602, 264603, 284604, 284605, 284486
781	80250049 (1561, 1562)				284905, 284907, 265010, 284600, 284601, 18108362, 18108374, 284556
782	8758529 (1563, 1564)	Novel Protein sim. GBank gij4155447 (AE001517) - proline/betaine transporter [Helicobacter pylori J99]		UNCLASSIFIED	284605
783	16410791 (1565, 1566)				285020
784	80051197 (1567, 1568)			UNCLASSIFIED	284635, 33657023, 29331828, 265017, 284585, 284568
785	56073541 (1569, 1570)	Novel Protein sim. GBank gij3451335 (AC005525) - F22162_1 [Homo sapiens]	Contains protein domain (PF00047) - Immunoglobulin domain	struct	35688052, 264604
786	20438842 (1571, 1572)	Novel Protein sim. GBank gij138748 [sp]P10905 [UGPA_ECOLI - SN-GLYCEROL-3-PHOSPHATE TRANSPORT SYSTEM PERMEASE PROTEIN UGPA]		transport	284603
787	80258364 (1573, 1574)			UNCLASSIFIED	284593

788	80507844 (1575, 1576)	Novel Protein sim. GBank gij2746079 (AF015310) - BTH1 [Brassica napus]		synthase	264909, 264602, 264603, 264769, 264638
789	17294715 (1577, 1578)	Novel Protein sim. GBank gij2351849 (U93357) - 40 kDa heat shock chaperone protein [Halobacterium salinarum]		UNCLASSIFIED	265007
790	86284408 (1579, 1580)	Novel Protein sim. GBank gij5706378 [dbj]BAA03099.1] - (AB026118) MAL T1 [Homo sapiens]	Contains protein domain (PF00047) - Immunoglobulin domain	glycoprotein	22278999, 29331824, 264828, 87188559, 265018, 21908765, 21908767, 21908768, 21908769, 265020, 264692, 22279000, 264563
791	84651827 (1581, 1582)	Novel Protein sim. GBank gij5689948 [emb]CAB51985.1] - (AL109863) putative isoleucyl-tRNA synthetase [Streptomyces coelicolor A3(2)]			264801, 264605, 264638
792	80056786 (1583, 1584)	Novel Protein sim. GBank gij393194 (L02375) - S-antigen [Plasmodium falciparum]		struct	265021, 264631, 264635, 264556
793	79636730 (1585, 1586)	Novel Protein sim. GBank gij1345408 [dbj]BAA05046] - (D26046) AT motif-binding factor [Mus musculus]	Contains protein domain (PF00046) - Homeobox domain	homeobox	264693
794	81839284 (1587, 1588)	Novel Protein sim. GBank gij105884 [pir]S24023 - dopamine receptor D4 - human (fragment)		UNCLASSIFIED	264603, 264604, 264910, 264762, 264906, 264638, 264909, 264757
795	80074988 (1589, 1590)	Novel Protein sim. GBank gij1877334 [emb]CAB07082] - (Z92771) bira [Mycobacterium tuberculosis]		carboxylase	264488, 35696052, 264905, 264907, 265010, 35698423, 264636
796	86669451 (1591, 1592)				80432229, 55811150, 264630, 264637, 264565
797	87771781 (1593, 1594)	Novel Protein sim. GBank gij2995447 [emb]CAA71519] - (Y10495) CDV-1R protein [Mus musculus]		struct	22278998, 264093, 264094, 66714117, 21908787, 21908769, 265020, 265022
798	79855209 (1595, 1596)			transcript factor	264687, 264768, 264693
799	79557816 (1597, 1598)	Novel Protein sim. GBank gij4467250 [emb]CAB37575] - (AL035569) probable Glu-tRNA Gln amidotransferase subunit [Streptomyces coelicolor]		hydrolase	264909, 264910, 264638, 264638
800	79970189 (1599, 1600)			UNCLASSIFIED	264488
801	80499399 (1601, 1602)	Novel Protein sim. GBank gij2791517 [emb]CAA16054] - (AL021246) hypothetical protein Rv2477c [Mycobacterium tuberculosis]		transport	264508, 264511, 265006, 265009, 264769, 264567, 264486
802	79834598 (1603, 1604)	Novel Protein sim. GBank gij488721 [gb]AAD32237.1 [AF14744] - (AF147449) penicillin binding protein 1B [Pseudomonas aeruginosa]			264905, 264693
803	20467520 (1605, 1606)			struct	264605
804	10174239 (1607, 1608)	Novel Protein sim. GBank gij1176152 [sp]P44507 [YHAD_HAEIN - HYPOTHETICAL PROTEIN H10091]		kinase	264510
805	79599993 (1609, 1610)				264508
806	80484113 (1611, 1612)	Novel Protein sim. GBank gij2764612 [emb]CAA04683] - (AJ001330) ornithine transcarbamoylase [Lactobacillus sakei]	Contains protein domain (PF00185) - Aspartate/ornithine carbamoyltransferase	transferase	264769
807	80361612 (1613, 1614)	Novel Protein sim. GBank gij283331 [sp]Q21828 [YNF0 CAEEI - HYPOTHETICAL 18.9 KD PROTEIN R07E5.13 IN CHR. II, SUME III]			264764

808	35106817 (1615, 1616)	Novel Protein sim. GBank gi 3913092 sp Q46170 ARCD_CLOPE - ARGININE/ORNITHINE ANTI-PORTER		transport	264909, 264602, 21908764, 18108374
809	81454254 (1617, 1618)	Novel Protein sim. GBank gi 3913010 sp P74309 ALF1_SYN3 - FRUCTOSE- BISPHOSPHATE ALDOLASE CLASS I (FBP ALDOLASE)	Contains protein domain (PF00274) - Fructose-bisphosphate aldolase class-I	UNCLASSIFIED	264508, 264906, 264909, 265007, 264910, 264758, 264600, 264602, 264603, 264605, 264687, 264769, 264689, 264636, 264488
810	80192761 (1619, 1620)	Novel Protein sim. GBank gi 401472 sp P30863 YAFB_ECOLI - HYPOTHETICAL OXIDOREDUCTASE IN ASPU-MLTD INTERGENIC REGION	Contains protein domain (PF00248) - Aldo-keto reductase family	reductase	284369
811	80079280 (1621, 1622)			UNCLASSIFIED	284558
812	10297654 (1623, 1624)			UNCLASSIFIED	264692
813	79612280 (1625, 1626)			UNCLASSIFIED	264906
814	80473427 (1627, 1628)	Novel Protein sim. GBank gi 146168 (J01617) - glutaminyl- (RNA synthetase [Escherichia coli])		synthase	264905, 264602, 264605, 264682, 264687, 264789, 264636
815	95419513 (1629, 1630)	Novel Protein sim. GBank gi 4589652 dbj BAA76848.1 - (AB023221) KIAA1004 protein [Homo sapiens]		UNCLASSIFIED	264488, 22278998, 22278999, 29331822, 29331824, 29331825, 29331827, 29331828, 29148499, 264905, 264908, 265007, 33657402, 60433358, 60433438, 264758, 265011, 265017, 265018, 265019, 264369, 264288, 264685, 21906765, 21906787, 265020, 265021, 264692, 65274620, 33657109, 264629, 18108376, 264635, 264638, 60170394, 58182323, 264564
816	18881910 (1631, 1632)				264600
817	95293318 (1633, 1634)	Novel Protein sim. GBank gi 1781144 emb CA806254 - (Z63868) hypothetical protein RV3069 [Mycobacterium tuberculosis]		UNCLASSIFIED	264595
818	90938190 (1635, 1636)	Novel Protein sim. GBank gi 1477468 (U35244) - vacuolar protein sorting homolog r-vps32a [Rattus norvegicus]			65274572, 22278999, 60424289, 35898052, 55812038, 21906768, 55811957, 35695917, 33657023, 18108370, 18108374, 55810764, 35896423, 55811576, 264636
819	80254977 (1637, 1638)	Novel Protein sim. GBank gi 1001352 dbj BAA10839 - (D84008) ABC transporter [Synecocystis sp.]		transport	264565
820	80059688 (1639, 1640)	Novel Protein sim. GBank gi 566914 sp P37484 YYBT_BACSU - HYPOTHETICAL 74.3 KD PROTEIN IN RPLI-COTF INTERGENIC REGION		UNCLASSIFIED	264600, 264602, 264604
821	79762590 (1641, 1642)			UNCLASSIFIED	264910
822	80215310 (1643, 1644)			UNCLASSIFIED	264510, 264594, 264637
823	94992299 (1645, 1646)	Novel Protein sim. GBank gi 3878400 emb CAA95828 - (Z71264) predicted using GeneFinder: Weak similarity to Mouse T-complex-associated-testes-expressed-1 protein (PIR Acc. No. A45841); cDNA EST EMBL:D32742 comes from this gene; cDNA EST EMBL:D33617 comes from this gene; cDNA EST...		struc	264509, 264687, 264691
824	80411171 (1647, 1648)	Novel Protein sim. GBank gi 1370076 emb CAA66887 - (X98235) type I [Drosophila melanogaster]	Contains protein domain (PF01429) - Methyl-CpG binding domain		264710, 264763, 264769, 264693

825	20638600 (1649, 1650)	Novel Protein sim. GBank gij3025132[sp]P77391[YEAG. ECOLI - HYPOTHETICAL 74.5 KD PROTEIN IN GAPA-RND INTERGENIC REGION (AL021648) hypothetical protein Rv3202c [Mycobacterium tuberculosis]		UNCLASSIFIED	284592
826	11075047 (1651, 1652)	Novel Protein sim. GBank gij3242281[emb]CAA16869] - (AL021648) hypothetical protein Rv3202c [Mycobacterium tuberculosis]			284605
827	80054207 (1653, 1654)	Novel Protein sim. GBank gij3417424[emb]CAA20312] - (AL031261) putative transport protein [Schizosaccharomyces pombe]			284603
828	95106322 (1655, 1656)	Novel Protein sim. GBank gij4336692[gb]AAD17697] - (AF101361) Abnormal X segregation [Drosophila melanogaster]	UNCLASSIFIED		52645080, 264508, 264905, 264509, 264908, 264907, 264908, 264909, 264512, 264910, 264591, 264758, 264600, 264768, 264768, 21908768, 35695917, 264691, 264628, 264630, 264631, 264632, 264634, 264635, 264636, 264637, 264556, 264639, 83373044, 18108385, 264563, 264566, 264488
829	81742215 (1657, 1658)	Novel Protein sim. GBank gij3020539 (AF080002) - UDP-N- acetyl(muramyl) tripeptide synthetase MurC [Helicobacter mobilis]	UNCLASSIFIED		284758, 264634
830	20396091 (1659, 1660)	Novel Protein sim. GBank gij1870004[emb]CAB08855] - (Z92539) hypothetical protein Rv1024 [Mycobacterium tuberculosis]	UNCLASSIFIED		284603
831	87112435 (1661, 1662)	Novel Protein sim. GBank gij2500056[sp]Q46267[PFLA_CLOPA - PYRUVATE FORMATE-LYASE ACTIVATING ENZYME	UNCLASSIFIED		86714117, 264910, 264639
832	19536322 (1663, 1664)	Novel Protein sim. GBank gij2497531[sp]Q46078[KPYK_CORGL - PYRUVATE KINASE (PK)	UNCLASSIFIED		284906
833	20726654 (1665, 1666)	Novel Protein sim. GBank gij2497531[sp]Q46078[KPYK_CORGL - PYRUVATE KINASE (PK)	UNCLASSIFIED		284602
834	21426762 (1667, 1668)	Novel Protein sim. GBank gij2497531[sp]Q46078[KPYK_CORGL - PYRUVATE KINASE (PK)	kinase		284600, 264602, 264768, 264689, 264636
835	84140482 (1669, 1670)	Novel Protein sim. GBank gij6989315 (U15184) - phosphate transport protein PSTC [Mycobacterium leprae]	Contains protein domain (PF00224) - Pyruvate kinase		264688, 263994, 21908767, 264910, 264632, 264635, 264259, 264639, 264693, 83373044, 264758, 35696052, 22279002, 264508, 264905, 264908, 264448, 263972, 264908, 264909
836	56126552 (1671, 1672)	Novel Protein sim. GBank gij2498533[sp]Q50598[Y0D8_MYCTU - HYPOTHETICAL 69.9 KD PROTEIN CY1A11.08	transport		35695917, 264557
837	79450450 (1673, 1674)	Novel Protein sim. GBank gij728867[sp]P40802[JAG_ARATH - ANTER-SPECIFIC PROLINE-RICH PROTEIN APC PRECURSOR	UNCLASSIFIED		264595
838	79184203 (1675, 1676)	Novel Protein sim. GBank gij728867[sp]P40802[JAG_ARATH - ANTER-SPECIFIC PROLINE-RICH PROTEIN APC PRECURSOR	UNCLASSIFIED		264687
839	79641125 (1677, 1678)	Novel Protein sim. GBank gij728867[sp]P40802[JAG_ARATH - ANTER-SPECIFIC PROLINE-RICH PROTEIN APC PRECURSOR	UNCLASSIFIED		264908
840	80056851 (1679, 1680)	Novel Protein sim. GBank gij4557753[re]INP_000372.1[pm]D1 - midline 1 protein finger)	Contains protein domain (PF000097) - Zinc finger, C3HC4 type (RING finger)		264762, 264556

841	80376318 (1881, 1882)	Novel Protein sim. GBank gij139805[sp]P08045[XFIN_XENLA - XFIN PROTEIN]	Contains protein domain (PF000096) - Zinc finger, C2H2 type	transcriptionfactor	284784
842	80078724 (1883, 1884)	Novel Protein sim. GBank gij2114321[dbj]BAA200371 - (D88733) membrane glycoprotein [Equine herpesvirus 1]	Contains protein domain (PF00569) - Zinc finger present in dystrophin, CBP/p300	UNCLASSIFIED	284905, 284908, 265008, 265009, 18108374, 56182323, 264558
843	87002847 (1885, 1886)	Novel Protein sim. GBank gij3882325[dbj]BAA34522.11 - (AB018345) KIAA0802 protein [Homo sapiens]	Contains protein domain (PF00170) - bZIP transcription factor	struct	284091, 28331825, 284806, 264768, 264583
844	17841439 (1887, 1888)	Novel Protein sim. GBank gij2224721[dbj]BAA20844 - (AB002388) KIAA0390 [Homo sapiens]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	transcriptionfactor	265011
845	18346844 (1889, 1890)				284629
846	79863441 (1891, 1892)	Novel Protein sim. GBank gij825679[pir]A36929 - virulence regulatory protein VarB - Pseudomonas solanacearum		kinase	284907
847	78695348 (1893, 1894)				264909
848	78489365 (1895, 1896)			UNCLASSIFIED	265020
849	79756367 (1897, 1898)			UNCLASSIFIED	264568
850	78817849 (1899, 1700)	Novel Protein sim. GBank gij183245[sp]P78061[CYJ_K_ECOLI - PUTATIVE GLUTAMINE SYNTHETASE (GLUTAMATE--AMMONIA LIGASE)]	Contains protein domain (PF00120) - Glutamine synthetase	UNCLASSIFIED	284809
851	95320333 (1701, 1702)	Novel Protein sim. GBank gij5454130[ref]NP_008280.1p[TLN] - talin	Contains protein domain (PF01608) - ILWEEQ domain		284488, 52844507, 284489, 18108398, 85274572, 56182575, 22278994, 22278995, 22278996, 35696286, 22278997, 22278998, 22278999, 20281171, 264490, 264259, 52845080, 29331822, 29331824, 66714117, 29331825, 60432289, 29331826, 29331827, 35896052, 29331828, 29146498, 29146499, 284107, 264905, 264906, 264907, 264908, 52844045, 56182435, 265008, 265007, 265008, 265009, 264910, 60432229, 60431735, 60433356, 33857402, 60433438, 284595, 264758, 284759, 21906754, 33109954, 52844296, 265010, 265011, 87168559, 285017, 285018, 265019, 284760, 264761, 264762, 264681, 18108351, 264763, 284448, 284682, 284784, 264683, 18108354, 284288, 284369, 264685, 264766, 284687, 284768, 284769, 21908765, 21908766, 21908767, 21908768, 29146627, 21906769, 29146628, 55811957, 35695917, 265020, 265021, 265022, 60170615, 52844150, 264691, 284692, 33857023, 284693, 283986, 33857109, 27486261, 27486262, 27486264, 27486265, 35895763, 60431602, 18108370, 20281089, 284629, 18108374, 18108376, 55811576, 35898423, 35895855, 284634, 284635, 264636, 264555, 60431850, 284556, 284691
852	10147368 (1703, 1704)				

853	13032587 (1705, 1706)	Novel Protein sim. GBank gij3402836[embjCAA76082] - (Y18136) 2-enoate reductase [Mooresella thermoacetic]		reductase	264638 264568
854	80052438 (1707, 1708)				
855	79641130 (1709, 1710)				264692
856	11594238 (1711, 1712)			UNCLASSIFIED	264591
857	79210165 (1713, 1714)			UNCLASSIFIED	264630, 264634
858	80246910 (1715, 1716)				265008, 265009, 264601, 264602, 264603, 18108351
859	20296634 (1717, 1718)				264559
860	80041749 (1718, 1720)			UNCLASSIFIED	264489
861	85857045 (1721, 1722)			UNCLASSIFIED	33857023, 264630
862	80079487 (1723, 1724)				264600
863	80578931 (1725, 1726)	Novel Protein sim. GBank gij2246532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus]		UNCLASSIFIED	264488, 18108398, 35696286, 264259, 18108351, 264288, 265021
864	94839904 (1727, 1728)			UNCLASSIFIED	264259, 264112, 263974
865	80045310 (1728, 1730)	Novel Protein sim. GBank gij5689884[embjCAB52047.1] - (AL109732) hypothetical protein [Streptomyces coelicolor A3(2)]	Contains protein domain (PF01479) - S4 domain		264635, 264600, 264636, 264591, 264602, 264693
866	80162031 (1731, 1732)	Novel Protein sim. GBank gij4557876[refjNP_000341.1]pABCR - ATP binding cassette transporter		transport	264288, 264557, 264558
867	80062402 (1733, 1734)				
868	10075384 (1735, 1736)			UNCLASSIFIED	264605
869	80062408 (1737, 1738)				264909
870	80249651 (1739, 1740)	Novel Protein sim. GBank gij828660[pirjIS37755 - Adenyl- transferase - Escherichia coli		transferase	264605, 264687, 18108374
871	20378295 (1741, 1742)	Novel Protein sim. GBank gij1708180[spjQ10602]HEMK_MYCTU - HEMK PROTEIN HOMOLOG		UNCLASSIFIED	264601, 264636
872	95197114 (1743, 1744)	Novel Protein sim. GBank gij1545959[embjCAA67763] - (X99384) paladin [Mus musculus]			264603
				UNCLASSIFIED	35686288, 22278998, 264259, 29331822, 29331824, 29331825, 60432289, 29331828, 29331827, 29331828, 35696052, 264509, 264905, 264908, 264907, 264908, 264909, 264510, 265008, 264511, 264512, 265007, 265008, 265009, 264910, 264591, 60433356, 264596, 52846317, 87168474, 265010, 264602, 264603, 265017, 265018, 264605, 18108351, 264764, 264768, 264768
					52844229, 264789, 21906765, 265021, 264534, 264691, 52645129, 264628, 264629, 35698423, 65274791, 264631, 264632, 264635, 264636, 264558, 264637, 264638, 264639, 60432113, 22279000, 22279002, 264584
873	20189728 (1745, 1746)	Novel Protein sim. GBank gij4156104 (AE001569) - putative Outer membrane protein [Helicobacter pylori J99]	Contains protein domain (PF00008) - Zinc finger, C2H2 type	UNCLASSIFIED	264595

874	80077682 (1747, 1748)	Novel Protein sim. GBank gll134319[sp]p07819[SCRB_BACSU - SUCROSE-8- PHOSPHATE HYDROLASE (SUCRASE) (INVERTASE)]		UNCLASSIFIED	264600
875	86608446 (1749, 1750)	Novel Protein sim. GBank gll481000[pi]q[S37594 - mucin - human (fragment)]			264259, 264448, 264288, 264557, 87165518
876	86465157 (1751, 1752)	Novel Protein sim. GBank gll3128283 (AF010496) - iron(ii) diclrate transport ATP-binding protein [Rhodobacter capsulatus]	Contains protein domain (PF00005) - transport ABC transporter		264907, 264601, 264602, 264605, 265020, 60431602
877	87802548 (1753, 1754)	Novel Protein sim. GBank gll731074[sp]p40349[URB1_USTMA - SIDEROPHORE BIOSYNTHESIS REGULATORY PROTEIN URBS1]	Contains protein domain (PF00320) - GATA zinc finger	transcriptfactor	22278998, 264908, 264369
878	80187289 (1755, 1756)	Novel Protein sim. GBank gll1351614[sp]q09853[VAED_SCHPO - HYPOTHETICAL 181.5 KD PROTEIN C23D3.13C IN CHROMOSOME I]		ATPase_associated	264369, 264555
878	94328962 (1757, 1758)	Novel Protein sim. GBank gll3875304[emb]CAA98434] - (Z74030) predicted using Genefinder; cDNA EST EMBL-C07608 comes from this gene; cDNA EST EMBL-C08023 comes from this gene; cDNA EST yk505e9.3 comes from this gene; cDNA EST yk489h9.3 comes from this gene; cDNA EST yk489h9.5 com...			56182575, 29331824, 264508, 264908, 265018, 18108351, 264448, 264883, 21908768, 21908768, 60170815, 33657023, 65274820, 33657109, 18108374, 35695855, 264583
880	8491135 (1759, 1760)	Novel Protein sim. GBank gll137120[sp]p11214[UR0T_MOUSE - TISSUE PLASMINOGEN ACTIVATOR PRECURSOR (TPA) (T- PLASMINOGEN ACTIVATOR)]	Contains protein domain (PF00051) - Kingle domain	cathepsin	264508
881	11280122 (1761, 1762)	Novel Protein sim. GBank gll2632098[emb]CAA75667] - (Y15513) Proctos protein [Drosophila melanogaster]		UNCLASSIFIED	264508
882	11077011 (1763, 1764)	Novel Protein sim. GBank gll1155068[emb]CAA64425] - (X94976) cell wall-plasma membrane linker protein [Brassica napus]		UNCLASSIFIED	264558
883	78582969 (1765, 1766)	Novel Protein sim. GBank gll1155068[emb]CAA64425] - (X94976) cell wall-plasma membrane linker protein [Brassica napus]		UNCLASSIFIED	264688
884	13517921 (1767, 1768)	Novel Protein sim. GBank gll2078027[emb]CAB08467] - (Z95208) hypothetical protein Rv2372c [Mycobacterium tuberculosis]		UNCLASSIFIED	264638
885	80052457 (1769, 1770)	Novel Protein sim. GBank gll2685834[emb]CAA15904] - (AL021006) sucA [Mycobacterium tuberculosis]		UNCLASSIFIED	264605, 18108362
886	11885138 (1771, 1772)	Novel Protein sim. GBank gll5689395[db]BAA82981.1] - (AB028952) KIAA1029 protein [Homo sapiens]			264690
887	94315307 (1773, 1774)	Novel Protein sim. GBank gll1181338[db]BAA19365] - (AB001488) PROBABLE INTEGRAL MEMBRANE PROTEIN, SIMILAR TO CHLORAMPHENICOL RESISTANCE PROTEIN OF STREPTOMYCES VENEZUELAE [Bacillus subtilis]		dehydrogenase	35698052, 264908, 264600, 264603, 35695917, 35685855, 264638
888	10083399 (1775, 1776)	Novel Protein sim. GBank gll1181338[db]BAA19365] - (AB001488) PROBABLE INTEGRAL MEMBRANE PROTEIN, SIMILAR TO CHLORAMPHENICOL RESISTANCE PROTEIN OF STREPTOMYCES VENEZUELAE [Bacillus subtilis]		UNCLASSIFIED	264908
889	20385917 (1777, 1778)	Novel Protein sim. GBank gll1181338[db]BAA19365] - (AB001488) PROBABLE INTEGRAL MEMBRANE PROTEIN, SIMILAR TO CHLORAMPHENICOL RESISTANCE PROTEIN OF STREPTOMYCES VENEZUELAE [Bacillus subtilis]			264603
890	19504337 (1779, 1780)	Novel Protein sim. GBank gll854085[emb]CAA58337] - (X83413) U88 [Human herpesvirus 6]			264629

891	13516879 (1781, 1782)	Novel Protein sim. GBank gi 458398 gb AAD34331.1 AF11248 - (AF112481) RAD54B protein [Homo sapiens]		UNCLASSIFIED	264638
892	87634157 (1763, 1784)	Novel Protein sim. GBank gi 545526 bb 143833 - LBP- 1b-transcription factor binding to initiation site of HIV-1 (alternatively spliced) [human, Namalwa cells, Peptide, 541 aa]		transcript factor	22278996, 22278999, 29331828, 35696052, 264908, 264909, 265009, 265011, 284602, 265019, 284786, 21908765, 21908788, 21908789, 265020, 265021, 58526488 264689, 263967
893	79168037 (1785, 1786)	Novel Protein sim. GBank gi 2829688 sp P80608 CYSK_MAIZE - CYSTEINE SYNTHASE (O-ACETYL SERINE SULFHYDRYLASE) (O- ACETYL SERINE (THIOL) LYASE) (CSASE)		synthase	
894	11102240 (1787, 1788)				263978
895	80239868 (1789, 1790)			UNCLASSIFIED	264508, 264600, 264555, 264559
896	79747803 (1791, 1792)				264632
897	94991823 (1793, 1794)			UNCLASSIFIED	264686, 29331828, 264511
898	67895109 (1795, 1796)				56182575, 60432289, 56182435, 60432229, 55811857, 22279000, 264488
899	11100463 (1797, 1798)				264601
900	80489788 (1798, 1800)	Novel Protein sim. GBank gi 1750127 (U66480) - YncC [Bacillus subtilis]		transport	264769, 264891, 264563
901	80502410 (1801, 1802)	Novel Protein sim. GBank gi 3122879 sp O07438 SYA_MYCTU - ALANYL-TRNA SYNTHETASE (ALANINE-TRNA LIGASE) (ALARS)		synthase	284907, 264602, 264605, 264769, 35695917, 18108376, 264583
902	80503301 (1803, 1804)	Novel Protein sim. GBank gi 335570 emb CAA20001 - (AL031124) 3-isopropylmalate dehydratase large subunit [Streptomyces coelicolor]		isomerase	264909, 265008, 264602, 264604, 264769, 264689, 264693
903	82060208 (1805, 1806)	Novel Protein sim. GBank gi 2960120 emb CAA18018.1 - (AL02:121) gipK [Mycobacterium tuberculosis]		kinase	35696052, 284905, 284510, 284511, 284512, 264605, 264760, 18108351, 264762, 284687, 264768, 264769, 264688, 21908784, 35695917, 27486262, 35695855, 284634, 284638, 264488
904	20451078 (1807, 1808)	Novel Protein sim. GBank gi 728687 sp P40908 ARGI_COCIM - ARGINASE		hydrolase	264604
905	9398483 (1808, 1810)	Novel Protein sim. GBank gi 4567200 gb AAD23616.1 AC00716 - (AC007168) hypothetical protein [Arabidopsis thaliana]		UNCLASSIFIED	264909
906	80052628 (1811, 1812)	Novel Protein sim. GBank gi 3738200 emb CAA21292 - (AL031855) putative vacuolar membrane protein [Schizosaccharomyces pombe]			264595, 264605
907	87913201 (1813, 1814)				
908	11754482 (1815, 1816)			UNCLASSIFIED	60432289, 264601, 264690
909	20727807 (1817, 1818)	Novel Protein sim. GBank gi 3668940 db BAA34296 - (AB015054) Alg2 [Rhizomucor pusillus]		UNCLASSIFIED	264638
910	16776206 (1819, 1820)	Novel Protein sim. GBank gi 4589728 db BAA76883.1 - (AB003137) DnaJ homolog protein [Salix glauca]		UNCLASSIFIED	264602
				eph	265009

911	87454340 (1821, 1822)	Novel Protein sim. GBank gij548774j sp P35685 RL7A_ORYSA - 60S RIBOSOMAL PROTEIN L7A		ribosomalprot	265010, 264604, 60432113
912	20448863 (1823, 1824)	Novel Protein sim. GBank gij2314008 gbi AAD07921.1 - (AE000597) CDP-diglyceride hydrolase (cdh) [Helicobacter pylori 26895]		hydrolase	264559
913	20469357 (1825, 1826)			UNCLASSIFIED	264604
914	79183351 (1827, 1828)	Novel Protein sim. GBank gij417657 sp Q03604 RIR1_CAEEL - PROBABLE RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE LARGE CHAIN (RIBONUCLEOTIDE REDUCTASE)	Contains protein domain (PF00317) - Ribonucleotide reductase	reductase	264636
915	87606703 (1829, 1830)	Novel Protein sim. GBank gij588957 dbj BAA83089.1 - (AB028040) KIAA1117 protein [Homo sapiens]			18108398, 22278996, 66714117, 264906, 264591, 21906788, 265020, 55811578, 264638
916	79444091 (1831, 1832)	Novel Protein sim. GBank gij4188110 emb CAA71790 - (Y10831) putative integrase [Ralstonia eutropha]		UNCLASSIFIED	264595
917	20185985 (1833, 1834)			UNCLASSIFIED	264605
918	91226795 (1835, 1836)	Novel Protein sim. GBank gij1655699 emb CAA69032 - (Y07752) pterophorin-S [Volvox carter]		synthase	264259, 26331826, 264908, 265019, 264448, 265020, 264635, 83373044
919	80438785 (1837, 1838)	Novel Protein sim. GBank gij5689668 emb CA852005.1 - (AL109663) putative membrane protein [Streptomyces coelicolor A3(2)]			265006, 264512, 264600, 264602, 264604, 264788, 18108370, 264583
920	78608095 (1839, 1840)	Novel Protein sim. GBank gij1168448 sp Q05813 AMP1_STRLI - XAA-PRO AMINOPEPTIDASE I (X-PRO AMINOPEPTIDASE I) (AMINOPEPTIDASE P I) (APP) (PEPP I) (AMINOACYLPROLINE AMINOPEPTIDASE I)		peptidase	264508
921	18658634 (1841, 1842)	Novel Protein sim. GBank gij3850084 emb CAA2181.1 - (AL033389) alcohol dehydrogenase [Schizosaccharomyces pombe]		UNCLASSIFIED	264600
922	78982605 (1843, 1844)				265019, 22279002
923	86659530 (1845, 1846)	Novel Protein sim. GBank gij207079 sp P29514 TB88_ARATH - TUBULIN BETA-6 CHAIN		tubulin	264907, 265008, 265009, 265010, 18108351, 264689, 265021, 18108370, 18108374, 18108385
924	21431341 (1847, 1848)				264510
925	20630332 (1849, 1850)	Novel Protein sim. GBank gij2497688 sp Q08983 pAFA_MOUSE - PLATELET- ACTIVATING FACTOR ACETYLHYDROLASE PRECURSOR (PAF ACETYLHYDROLASE) (PAF 2- ACYLHYDROLASE) (LDL-ASSOCIATED PHOSPHOLIPASE A2) (LDL-PLA2) (2-ACETYL-1- ALKYLGLYCEROPHOSPHOCHOLINE ESTERASE) (1- ALKYL-2-ACETYLGLYCEROPHO...		esterase	264603
926	79397657 (1851, 1852)	Novel Protein sim. GBank gij3882325 dbj BAA34522.1 - (AB018345) KIAA0802 protein [Homo sapiens]			55811957, 263972, 264639
927	37036201 (1853, 1854)			UNCLASSIFIED	264769

928	80070610 (1855, 1856)	Novel Protein sim. GBank gij158146 (M30316) - xanthine dehydrogenase (AA at 2538) [Calliphora vicina]		dehydrogenase	264605
929	20630336 (1857, 1858)			UNCLASSIFIED	264603
930	5496348 (1859, 1860)	Novel Protein sim. GBank gij4115938 [gbjAAD03446.1] - (AF118223) No definition line found [Arabidopsis thaliana]			264259
931	10245731 (1861, 1862)	Novel Protein sim. GBank gij4490608 [embjCAB38642.1] - (AJ133495) ribonucleotide reductase major subunit [Staphylococcus aureus]	Contains protein domain (PF00317) - Ribonucleotide reductase	reductase	264486
932	80420613 (1863, 1864)	Novel Protein sim. GBank gij5459396 [embjCAB50754.1] - (AL096839) putative integral membrane transport protein [Streptomyces coelicolor]		transport	264508, 264905, 264908, 264909, 264600, 264602, 264603, 264805, 264768, 264558, 18108387, 264486
933	94326010 (1865, 1866)	Novel Protein sim. GBank gij5699523 [dbjBAA83045.1] - (AB029016) KIAA1093 protein [Homo sapiens]	Contains protein domain (PF00827) - UBA domain	UNCLASSIFIED	264508, 264686, 264693, 27486281, 18108370, 65274781, 264636, 264559, 22279002
934	80039105 (1867, 1868)	Novel Protein sim. GBank gij11911 [spP12978] EBN2_EBV - EBNA-2 NUCLEAR PROTEIN			264369, 22279002
935	80063162 (1869, 1870)				
936	80026632 (1871, 1872)	Novel Protein sim. GBank gij845688 (M32103) - ORF-27 [Staphylococcus aureus]		UNCLASSIFIED	264604, 264605, 264693, 18108370, 18108374
937	80250273 (1873, 1874)	Novel Protein sim. GBank gij1360668 [pirjICGHU1V - collagen alpha 1(V) chain precursor - human]		UNCLASSIFIED	22278995, 22278996, 264602, 264687, 32833986, 18108387
938	80026633 (1875, 1876)	Novel Protein sim. GBank gij2262624 [embjCAAT4531.1] - (Y14083) hypothetical protein [Bacillus subtilis]		synthase	263978
939	11071694 (1877, 1878)			UNCLASSIFIED	264602
940	94144252 (1878, 1880)	Novel Protein sim. GBank gij3560166 [embjCAA20878] - (AL031525) ubiquitin carboxyl-terminal hydrolase [Schizosaccharomyces pombe]	Contains protein domain (PF00442) - Ubiquitin carboxyl-terminal hydrolases family 2	ubiquitin	264600
941	11398414 (1881, 1882)				264905, 264908, 264907, 264908, 264909, 264511, 264910, 264592, 33857402, 264596, 264758, 264760, 264683, 264768, 264768, 264769, 33657023, 33657109, 264628, 264629, 264630, 264635, 264638, 264555, 264637, 264556, 264638, 264639, 83373044, 18108385, 264565, 18108391
942	19484122 (1883, 1884)			UNCLASSIFIED	264593
943	80080258 (1885, 1886)	Novel Protein sim. GBank gij4033729 (AF038595) - apolipoprotein N-acyltransferase [Pseudomonas aeruginosa]		UNCLASSIFIED	264760
944	80218096 (1887, 1888)	Novel Protein sim. GBank gij2494784 [spIQ50728] GUAA_MYCTU - GMP SYNTHASE (GLUTAMINE-HYDROLYZING) (GLUTAMINE AMIDOTRANSFERASE) (GMP SYNTHETASE)		synthase	264600, 264687, 264688, 264563
945	80052477 (1889, 1890)	Novel Protein sim. GBank gij732353 [spP39606] YWCH_BACSU - HYPOTHETICAL 38.6 KD PROTEIN IN QOXD-YPR INTERGENIC REGION	Contains protein domain (PF00818) - Ice nucleation protein repeat		264511, 264603
946	70248402 (1891, 1892)			UNCLASSIFIED	264908, 264604, 264605, 265020, 18108387
					265017

947	81802699 (1893, 1894)	Novel Protein sim. GBank gl 2898770 emb CAA17247 - (AL021899) hypothetical protein Rv2033c [Mycobacterium tuberculosis]	Contains protein domain (PF000459) - Inositol monophosphatase family	phosphatase	18108394, 22278896, 264907, 264908, 285008, 265009, 264910, 264758, 264600, 264602, 265018, 264605, 264769, 264689, 264693
948	88165538 (1895, 1896)	Novel Protein sim. GBank gl 2827284 (AF041037) - novel antagonist of EGF signaling [Homo sapiens]		igl	18108398, 58182575, 22278897, 22278999, 60432049, 29331822, 29331826, 264907, 56182435, 55811388, 265011, 264600, 265017, 265018, 265019, 18108351, 265020, 265021, 265022, 27486265, 263972, 55811576, 264638, 80170394, 264568
949	88081786 (1897, 1898)	Novel Protein sim. GBank gl 4507985 ref NP_003427.1 pZNF1 - zinc finger protein 135 (clon pHZ-17)	Contains protein domain (PF00096) - Zinc finger, C2H2 type	transcriptfactor	29331825, 21908764, 27486261, 21908768, 52844298, 33657349, 87188518, 56994075, 265020, 265021, 87168559, 52644150, 264637
950	78485872 (1899, 1900)	Novel Protein sim. GBank gl 1079461 p S3865 - Cytochrome b, type II - polio (fragment)	Contains protein domain (PF00038) - Intermediate filament proteins	struct	264683, 18108361
951	20451411 (1901, 1902)	Novel Protein sim. GBank gl 5420387 emb CAB46879.1 - (AJ243459) proteophosphoglycan [Leishmania major]		UNCLASSIFIED	264604
952	78566954 (1903, 1904)	Novel Protein sim. GBank gl 535702 gb AADA1779.1 AF12686 - (AF126867) calpain-like protease [Mus musculus]		calhepsin	264910, 264691
953	10196003 (1905, 1906)	Novel Protein sim. GBank gl 2495842 sp Q47142 YFHS_ECOLI - HYPOTHETICAL 41.8 KD PROTEIN IN CSIE-GLYA INTERGENIC REGION		transport	264510
954	9893326 (1907, 1908)	Novel Protein sim. GBank gl 2360965 (AF016253) - D-amino acid dehydrogenase [Klebsiella aerogenes]		dehydrogenase	264508
955	95313410 (1909, 1910)	Novel Protein sim. GBank gl 5454064 ref NP_006319.1 pSIP1 - SYT Interacting protein	Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	dna_rna_bind	56994075, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 264910, 264758, 264759, 265010, 264601, 264760, 18108351, 264762, 264763, 264764, 264766, 264686, 264767, 264687, 264768, 264769, 264689, 264628, 264629, 264630, 264631, 264632, 264634, 264635, 264636, 264637, 264638, 56182323, 264639, 18108388, 264583, 264564
956	80064224 (1911, 1912)	Novel Protein sim. GBank gl 2052128 emb CAB08155 - (Z94752) rimJ [Mycobacterium tuberculosis]			264605
957	80056208 (1913, 1914)	Novel Protein sim. GBank gl 1709787 sp Q00451 PRF1_LYCES - 36.4 KD PROLINE-RICH PROTEIN		UNCLASSIFIED	264603, 18108362
958	80036448 (1915, 1916)	Novel Protein sim. GBank gl 1709787 sp Q00451 PRF1_LYCES - 36.4 KD PROLINE-RICH PROTEIN		UNCLASSIFIED	264908, 264910, 264762, 263978, 264637
959	80026847 (1917, 1918)	Novel Protein sim. GBank gl 2131050 emb CAB09260 - (Z95844) opcA [Mycobacterium tuberculosis]		UNCLASSIFIED	264602, 264692
960	37815406 (1919, 1920)	Novel Protein sim. GBank gl 2129478 p S1939 - chitinase (EC 3.2.1.14) precursor - beet		UNCLASSIFIED	264259
961	20587383 (1921, 1922)				263978
962	11399318 (1923, 1924)				264593

963	80580374 (1925, 1926)			UNCLASSIFIED	284510, 284288, 284555, 284556, 284559, 284488
964	79832019 (1927, 1928)	Novel Protein sim. GBank gi 4588822 dbj BAA78833.1 - (AB023206) KIAA0989 protein [Homo sapiens]		UNCLASSIFIED	284112, 284910, 284689
965	81228485 (1929, 1930)	Novel Protein sim. GBank gi 5420387 emb CAB46679.1 - (AJ243459) proteophosphoglycan [Leishmania major]		UNCLASSIFIED	284488, 285017, 284440, 284634, 284558, 83373044
966	95282815 (1931, 1932)			UNCLASSIFIED	284908, 284592, 284596, 284604, 284788, 21908764, 284892, 284893, 284829, 284638, 284639
967	79255708 (1933, 1934)	Novel Protein sim. GBank gi 1731207 sp Q11156 RCX3_MYCTU - SENSORY TRANSDUCTION PROTEIN REGX3	Contains protein domain (PF00072) - Response regulator receiver domain	phosphatase	284760
968	79560268 (1935, 1936)	Novel Protein sim. GBank gi 2661836 emb CAA751871.1- (Y14984) putative transport protein [Methylophilus methylotrophus]		transport	284693
969	79919470 (1937, 1938)	Novel Protein sim. GBank gi 5418878 emb CAB46422.1 - (AL098747) hypothetical protein [Homo sapiens]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	dna_ma_bind	35698286, 284885, 284888, 284889, 35959517, 284892, 18108374, 284835
970	95085947 (1939, 1940)			UNCLASSIFIED	18108392, 18108394, 18108398, 22278995, 22278996, 22278998, 22278999, 28147820, 284828, 285006, 285007, 285008, 285009, 18108346, 33108954, 285010, 285011, 18108351, 284288, 21908767, 21908768, 18108370, 18108374, 18108377, 284630, 284635, 18108380, 83373044, 18108387, 18108388
971	78919770 (1941, 1942)			UNCLASSIFIED	285007, 285020, 22278002
972	20710704 (1943, 1944)				284557
973	20370183 (1945, 1946)	Novel Protein sim. GBank gi 1723119 sp P53990 Y174_HUMAN - HYPOTHETICAL PROTEIN KIAA0174			284604
974	80057103 (1947, 1948)			UNCLASSIFIED	284555
975	10186018 (1949, 1950)			UNCLASSIFIED	284510
976	80205742 (1951, 1952)	Novel Protein sim. GBank gi 3881459 emb CAA92988.1 - (Z68753) predicted using Genefinder; Similarity to Yeast hypothetical protein YIK9 (SW:YIK9_YEAST); cDNA EST EMBL:D27680 comes from this gene; cDNA EST EMBL:D27679 comes from this gene; cDNA EST EMBL:D84477 comes from this ge...		UNCLASSIFIED	284508, 284908, 284758, 284632, 284639, 284563
977	10355349 (1953, 1954)	Novel Protein sim. GBank gi 549456 sp Q05335 XYS3_PSEPU - XYLDLEGF OPERON TRANSCRIPTIONAL ACTIVATOR 3		UNCLASSIFIED	284906
978	80025927 (1955, 1956)			UNCLASSIFIED	284600, 284602, 284603, 284604
979	80447820 (1957, 1958)	Novel Protein sim. GBank gi 3171904 emb CAA75669.1- (Y15908) DIA-12C protein [Homo sapiens]		UNCLASSIFIED	284767, 284768, 285008, 285007, 284906
980	80025928 (1959, 1960)			UNCLASSIFIED	
981	80098550 (1961, 1962)	Novel Protein sim. GBank gi 3599940 (AF017368) - facogenital dysplasia protein 2 [Mus musculus]		UNCLASSIFIED	284600, 284602, 284605
				UNCLASSIFIED	284692, 284555, 284556, 284557, 284559

982	80195670 (1983, 1984)	Novel Protein sim. GBank gl 2850220 emb CAA71575 - (Y10545) fused-ccdB [Escherichia coli]		UNCLASSIFIED	264404
983	90995041 (1965, 1966)	Novel Protein sim. GBank gl 476389 pir B43402 - myosin heavy chain-B, neuronal - chicken		struct	65274572, 56182575, 264908, 264909, 265007, 265008, 264758, 265010, 55811150, 33857023, 284634, 284557, 284558
984	20466378 (1987, 1988)	Novel Protein sim. GBank gl 3451504 emb CAA07660.1 - (AJ007747) hypothetical protein BbLPS1.21 [Bordetella bronchiseptica]	Contains protein domain (PF00534) - Glycosyl transferases group 1	UNCLASSIFIED	264605
985	55461368 (1969, 1970)	Novel Protein sim. GBank gl 3451504 emb CAA07660.1 - (AJ007747) hypothetical protein BbLPS1.21 [Bordetella bronchiseptica]		transferase	56182435, 284600
986	87102888 (1971, 1972)			UNCLASSIFIED	264106, 264110, 265020, 60170615
987	79867231 (1973, 1974)			UNCLASSIFIED	264909
988	19858661 (1975, 1976)			UNCLASSIFIED	264800
989	88095329 (1977, 1978)	Novel Protein sim. GBank gl 3725508 gb AAD48080.1 AF08015 - (AF060152) METH1 protein [Homo sapiens]	Contains protein domain (PF01421) - Reprolysin (M12B) family zinc metalloprotease	UNCLASSIFIED	264508, 265017, 284534, 284584
990	88057746 (1978, 1980)	Novel Protein sim. GBank gl 2105049 emb CAB08835 - (Z85436) hypothetical protein Rv3645 [Mycobacterium tuberculosis]	Adenylate and Guanylate cyclase catalytic domain	UNCLASSIFIED	264259, 284908, 285009, 284910, 284598, 264369, 284288, 264768, 284628, 284635, 264568
991	10106140 (1981, 1982)	Novel Protein sim. GBank gl 3329297 AE001355 - Ribonucleoside Reductase, Large Chain [Chlamydia trachomatis]		UNCLASSIFIED	284909
992	79845694 (1983, 1984)	Novel Protein sim. GBank gl 3329297 AE001355 - Ribonucleoside Reductase, Large Chain [Chlamydia trachomatis]		UNCLASSIFIED	284508, 284593
993	10814053 (1985, 1986)	Novel Protein sim. GBank gl 5106572 gb AAD39760.1 AF14394 - (AF143946) transcriptional activator SRCAP [Homo sapiens]		reductase	264907
994	11090390 (1987, 1988)	Novel Protein sim. GBank gl 5106572 gb AAD39760.1 AF14394 - (AF143946) transcriptional activator SRCAP [Homo sapiens]	Contains protein domain (PF00176) - SNF2 and others N-terminal domain	helicase	264602
995	94321911 (1988, 1990)	Novel Protein sim. GBank gl 5106572 gb AAD39760.1 AF14394 - (AF143946) transcriptional activator SRCAP [Homo sapiens]			18108398, 65274572, 22278996, 264490, 80432049, 29331827, 28148498, 284508, 264905, 264907, 284908, 56182435, 265008, 264591, 264592, 60432228, 60431735, 33657402, 284595, 264758, 21906754, 265010, 265017, 265018, 264605, 284760, 284448, 284763, 284768, 21908785, 21908768, 21908769, 55811957, 264692, 264893, 264629, 35696423, 55811576, 35695855, 284638, 284555, 264558, 264558, 83373044, 22278002, 264583
996	91013745 (1991, 1992)	Novel Protein sim. GBank gl 2811719 AC004227 - KIA0811B [Homo sapiens]	Contains protein domain (PF00585) - PDZ domain (Also known as DHR or GLGF)	phosphatase	83373044, 29331824, 263978, 55811957, 56526488, 87168518, 284910, 284908, 264585, 264586, 264693, 284768
997	80503347 (1993, 1994)	Novel Protein sim. GBank gl 2649101 AE001001 - ABC transporter, ATP-binding protein [Archaeoglobus fulgidus]	Contains protein domain (PF00005) - ABC transporter	transport	35696286, 22278997, 22278999, 264508, 264905, 264908, 265010, 264600, 264602, 264605, 284688, 284769, 265021, 284585, 18108391
998	11397380 (1995, 1996)	Novel Protein sim. GBank gl 123530 sp P04929 HRPX_PLALO - HISTIDINE-RICH GLYCOPROTEIN PRECURSOR			284595

999	11768047 (1987, 1998)	Novel Protein sim. GBank gi 2508697 sp P46480 YFCA_HAEIN - HYPOTHETICAL PROTEIN H10198		UNCLASSIFIED	284682	
1000	20727944 (1999, 2000)			UNCLASSIFIED	284602	
1001	86673131 (2001, 2002)	Novel Protein sim. GBank gi 2224699 dbj BAA20833 - (AB002377) KIAA0379 [Homo sapiens]	Contains protein domain (PF00023) - Ank repeat	UNCLASSIFIED -kinase	60432049, 284907, 284908, 284511, 284603, 284683, 284684, 284687, 284689, 29148827, 21908769, 284692, 18108385, 22279000 285009, 284369, 285020	
1002	80189603 (2003, 2004)	Novel Protein sim. GBank gi 588121 sp P37709 TRHY_RABIT - TRICHOHYALIN		struct		
1003	17933491 (2005, 2006)				285019	
1004	16314887 (2007, 2008)	Novel Protein sim. GBank gi 854065 emb CAA58337 - (X83413) U88 [Human herpesvirus 6]			284635	
1005	79617144 (2009, 2010)	Novel Protein sim. GBank gi 114073 sp P07872 APT_ECOLI - ADENINE PHOSPHORIBOSYLTRANSFERASE (APRT)		UNCLASSIFIED	284508	
1006	37815429 (2011, 2012)					
1007	79620871 (2013, 2014)	Novel Protein sim. GBank gi 4062979 dbj BAA36210.1 - (AB017138) epsilon subunit of malonate decarboxylase [Pseudomonas putida]		UNCLASSIFIED synthase	284239 284805	
1008	86094444 (2015, 2016)	Novel Protein sim. GBank gi 2808807 emb CAA04607.1 - (AJ001208) putative trehalose synthase [Streptomyces coelicolor]		synthase	285007, 284802, 284605, 284760, 284638	
1009	57451289 (2017, 2018)	Novel Protein sim. GBank gi 3639077 (AF080113) - AMPA receptor binding protein [Rattus norvegicus]	Contains protein domain (PF00595) - PDZ domain (Also known as DHR or GLGF)	kinase	284102, 284288	
1010	94672537 (2019, 2020)	Novel Protein sim. GBank gi 3746332 (AF016307) - possible NADH-dependent oxidase, may function as a demethylase [Sinorhizobium meliloti]		dehydrogenase	284592	
1011	85546918 (2021, 2022)	Novel Protein sim. GBank gi 2342847 gb AAB86591.1 - (U90653) DHC-domain-containing cysteine-rich protein [Homo sapiens]		UNCLASSIFIED	35696052, 284905, 284764, 284788, 35895917, 284629	
1012	95294456 (2023, 2024)	Novel Protein sim. GBank gi 3413411 emb CAA20272 - (AL031231) guanosine pentaphosphate synthetase/ polyribonucleotide nucleotidyltransferase [Streptomyces coelicolor]	Contains protein domain (PF00013) - KH domain	phosphorylase	35696052, 284905, 284600, 284601, 284602, 284605, 284782, 284786, 284788, 284689	
1013	86095772 (2025, 2026)			UNCLASSIFIED	284591, 21908788	
1014	86608828 (2027, 2028)				29331824, 285019, 285020	

1015	95418679 (2029, 2030)	Novel Protein sim. GBank gl[4159995 (AF063095) - SELIL [Mus musculus]	Contains protein domain (PF00040) - Fibronectin type II domain	struct	22278994, 22278995, 56994075, 22278998, 22278999, 284259, 29331825, 29331826, 264907, 56182435, 264510, 264591, 264593, 60433358, 264594, 55812038, 264758, 21908754, 33657084, 265010, 264600, 265017, 265018, 265019, 18108351, 21906765, 21906766, 21906767, 21906768, 55811957, 265022, 33657023, 65274620, 33657182, 32833986, 18108370, 18108377, 55811576, 35898423, 264630, 22279000, 284565 264686, 264693
1016	78559694 (2031, 2032)	Novel Protein sim. GBank gl[25069694]p41407IACPD_ECOLI - ACYL CARRIER PROTEIN PHOSPHODIESTERASE (ACP PHOSPHODIESTERASE)		esterase	
1017	11069213 (2033, 2034)	Novel Protein sim. GBank gl[5103943]dbj[BAA79259.1] - (AP000059) 802aa long hypothetical oligopeptide-binding protein oppA [Acetopyrum pernix]	Contains protein domain (PF00498) - Bacterial extracellular solute-binding proteins, family 5	transport	264600
1018	80072430 (2035, 2036)	Novel Protein sim. GBank gl[4493973]embj[CAB39032.1] - (AL034559) predicted using hexExon; MAL3P7.14 (PFC0925w). Hypothetical protein, len: 489 aa [Plasmodium falciparum]			22278996, 29148627, 264563
1019	11703607 (2037, 2038)			UNCLASSIFIED	264686
1020	80234432 (2039, 2040)				264508, 264509, 284312, 264600, 284762, 264769, 264689, 18108370, 264636, 264638, 264488
1021	37036243 (2041, 2042)	Novel Protein sim. GBank gl[4633607]gbjAAD26859.1AF12779 - (AF127795) trehalose biosynthetic enzyme TreY [Rhizobium leguminosarum bv. viciae]		synthase	264769
1022	80502627 (2043, 2044)	Novel Protein sim. GBank gl[1781230]embj[CAB06277] - (Z83867) hypothetical protein Rv3137 [Mycobacterium tuberculosis]	Contains protein domain (PF00459) - Inositol monophosphatase family	phosphatase	35686052, 264508, 265008, 265009, 264769, 18108387, 284563
1023	11399341 (2045, 2046)	Novel Protein sim. GBank gl[3777495 (U92083) - calcium transporting ATPase [Pichia angusta]	Contains protein domain (PF00122) - E1-E2 ATPase	ATPase_associated	264593
1024	80057129 (2047, 2048)			UNCLASSIFIED	52846842, 33657402, 33657023, 18108379, 55811576, 264631, 264556, 264557, 264559, 18108386, 284566
1025	79644200 (2049, 2050)	Novel Protein sim. GBank gl[3483045]embj[CAB02556] - (AL031371) putative transport system permease protein [Streptomyces coelicolor]		transport	264693
1026	80025946 (2051, 2052)	Novel Protein sim. GBank gl[1749225]Q02322[UVRD_HAEIN - DNA HELICASE II		helicase	264602
1027	17659234 (2053, 2054)	Novel Protein sim. GBank gl[4757728]refNP_004886.1pAGTA - angiotensin vasopressin receptor AII/AVP-like		UNCLASSIFIED	265017

1028	20287928 (2055, 2056)	Novel Protein sim. GBank gij2791409[embjCAA16003] - (AL021184) acn [Mycobacterium tuberculosis]	Contains protein domain (PF00330) - Aconitase family (aconitate hydratase)	UNCLASSIFIED	284600
1029	94665080 (2057, 2058)			UNCLASSIFIED	284595
1030	86095343 (2059, 2060)			UNCLASSIFIED	284907, 284908, 284510, 284512, 285009, 285010, 285011, 284600, 284602, 284603, 284605, 284769, 18108372, 18108374
1031	95289117 (2061, 2062)			UNCLASSIFIED	284905, 284908, 284909, 284595, 284692, 284630, 284634, 284638
1032	94873275 (2063, 2064)	Novel Protein sim. GBank gij4503895[re]NP_000145.1[pGALK - galactokinase 1]		UNCLASSIFIED	284689
1033	86464818 (2065, 2066)	Novel Protein sim. GBank gij2982990 (AE000682) - hypothetical protein [Aquifex aeolicus]		UNCLASSIFIED	35896052, 284908, 284510, 18108354, 284687, 284769, 284689, 60431802, 18108385, 284486
1034	78245937 (2067, 2068)	Novel Protein sim. GBank gij405895 (U00007) - methionyl-tRNA synthetase [Escherichia coli]		UNCLASSIFIED	284906
1035	79956355 (2069, 2070)			UNCLASSIFIED	284892
1036	85804998 (2071, 2072)			UNCLASSIFIED	284905, 66712502, 284908, 284768
1037	87896038 (2073, 2074)			UNCLASSIFIED	29331824, 284909, 60433438, 285019
1038	20481015 (2075, 2076)	Novel Protein sim. GBank gij790819 (L39881) - polycystic kidney disease-associated protein [Homo sapiens]	Contains protein domain (PF01477) - PLATLH2 domain		284604, 284634
1039	87260021 (2077, 2078)	Novel Protein sim. GBank gij2605967 (AF030027) - 24 [Equine herpesvirus 4]		UNCLASSIFIED	284902, 284093, 284094, 284683, 284689, 283987
1040	80026840 (2079, 2080)	Novel Protein sim. GBank gij2352095 (U97022) - DNA topoisomerase I [Fervidobacterium islandicum]	Contains protein domain (PF01131) - Prokaryotic DNA topoisomerase	isomerase	284595
1041	10156682 (2081, 2082)	Novel Protein sim. GBank gij2256555[dbj]BAA29218.1 - (AP000001) 301aa long hypothetical 2-phosphoglycerate kinase [Pyrococcus horikoshii]		kinase	284807
1042	11084375 (2083, 2084)	Novel Protein sim. GBank gij2058299[embj]CAA66953 - (X98309) ARI protein [Drosophila melanogaster]			284605
1043	80057136 (2085, 2086)	Novel Protein sim. GBank gij1870167[embj]CAA70125 - (Y08921) msK [Streptomyces reticuli]	Contains protein domain (PF00005) - ABC transporter	transport	284565, 284587
1044	80025952 (2087, 2088)			UNCLASSIFIED	285008, 284602, 285017
1045	52415482 (2089, 2090)	Novel Protein sim. GBank gij5688950[embj]CAB52053.1 - (AL109732) hypothetical protein [Streptomyces coelicolor A3(2)]			29331825, 284637
1046	11754882 (2091, 2092)	Novel Protein sim. GBank gij854085[embj]CAA58337 - (X83413) U88 [human herpesvirus 6]		helicase	284686
1047	37036258 (2093, 2094)	Novel Protein sim. GBank gij4210471[dbj]BAA74535.1 - (AB019033) oriSA [Pseudomonas sp.]		UNCLASSIFIED	284769
1048	79188400 (2095, 2096)	Novel Protein sim. GBank gij3413419[embj]CAA20278 - (AL031232) hypothetical protein SC10H5.07 [Streptomyces coelicolor]	Contains protein domain (PF00220) - Neurohypophyseal hormones, N-terminal Domain	UNCLASSIFIED	284687
1049	81755108 (2097, 2098)	Novel Protein sim. GBank gij5051630[dbj]AAD38328.1[AF07372] - (AF07372) EH domain-binding mitotic phosphoprotein [Homo sapiens]		UNCLASSIFIED	284805, 284634
1050	78471521 (2099, 2100)			UNCLASSIFIED	284686

1051	80475471 (2101, 2102)			UNCLASSIFIED	18108374, 264769, 265010, 265011, 264601, 265009, 264604, 264605, 264636, 18108351, 264692
1052	8242862 (2103, 2104)	Novel Protein sim. GBank gi 3123275 p P35136 SERA_BACSU - D-3- PHOSPHOGLYCERATE DEHYDROGENASE (PGDH)	Contains protein domain (PF00389) - D-isomer specific 2-hydroxyacid dehydrogenases	dehydrogenase	264508, 264762, 264687, 264486
1053	94851640 (2105, 2106)	Novel Protein sim. GBank gi 5441318 emb CAB46717.1 - (AL034396) dJ1158B12.1 (zinc finger, X-linked, duplicated A) [Homo sapiens]			264688, 18108374, 26331824, 83373044, 21908754, 52845158, 56182435, 264689, 29331827, 27486281, 35698052, 21908765, 35698423, 21908768, 56182575, 21908769, 55811957, 87168518, 35698288, 22278997, 265020, 265011, 265021, 265022, 265007, 265018, 22279000, 22279002, 264482, 264906, 52844150, 264909, 264288, 29331822, 52845080, 264768
1054	78580225 (2107, 2108)			UNCLASSIFIED	264688
1055	80594138 (2109, 2110)	Novel Protein sim. GBank gi 5052508 gb AAD38594.1 AF14560 - (AF145609) BcDNA GH02833 [Drosophila melanogaster]	Contains protein domain (PF00270) - DEAD/DEAH box helicase	helicase	264907, 264602, 264681, 264288, 21906768, 33657109, 55810764, 35695855, 264631
1056	17882319 (2111, 2112)	Novel Protein sim. GBank gi 3021676 dbj BAA25358 - (D86033) RNA polymerase sigma-70 factor [Pseudomonas fluorescens]		mapolymerase	264906
1057	85667216 (2113, 2114)	Novel Protein sim. GBank gi 226281 U50308 - No definition line found [Caenorhabditis elegans]		UNCLASSIFIED	264682
1058	80376576 (2115, 2116)				264764
1059	84662754 (2117, 2118)	Novel Protein sim. GBank gi 1170016 sp p46808 GREY_MYCLE - TRANSCRIPTION ELONGATION FACTOR GREY (TRANSCRIPT CLEAVAGE FACTOR GREY)		transcriptfactor	35696032, 35695855, 265009, 264638
1060	79481169 (2119, 2120)	Novel Protein sim. GBank gi 2499087 sp Q09332 UGG3_DROME - UDP- GLUCOSE:GLYCOPROTEIN GLUCOSYLTRANSFERASE PRECURSOR (DUGT)		glycoprotein	29146499, 264681, 264683, 264687
1061	11034025 (2121, 2122)	Novel Protein sim. GBank gi 90254 pir J28334 - protein- lysine-phosphatase (EC 3.1.3.48) Ly-5 precursor (B-cell variant) - mouse		phosphatase	264634
1062	39567837 (2123, 2124)	Novel Protein sim. GBank gi 3334200 sp Q49954 GCSP_SOLTU - GLYCINE DEHYDROGENASE (DECARBOXYLATING) PRECURSOR (GLYCINE DECARBOXYLASE) (GLYCINE CLEAVAGE SYSTEM P-PROTEIN)		dehydrogenase	264593
1063	8490481 (2125, 2126)	Novel Protein sim. GBank gi 2499986 sp Q41228 PSE1_NICSY - PHOTOSYSTEM I REACTION CENTRE SUBUNIT IV A PRECURSOR (PSI-E A)			264508
1064	78891783 (2127, 2128)	Novel Protein sim. GBank gi 92654 pir JJA0086 - 10K zein precursor - maize			265007, 265008, 18108351, 18108385

1065	80021208 (2128, 2130)	Novel Protein sim. GBank gi 2120998 pir J S70682 - glycosyltransferase homolog - Bordetella pertussis		transferase	264600, 264602, 264689
1066	17896879 (2131, 2132)	Novel Protein sim. GBank gi 2506382 sp P15042 DNLJ_ECOLI - DNA LIGASE (POLYDEOXYRIBONUCLEOTIDE SYNTHASE (NAD+))		synthase	265009
1067	10132178 (2133, 2134)	Novel Protein sim. GBank gi 4007669 emb CAA22355 - (AL034443) putative oxidoreductase (Streptomyces coelicolor)	Contains protein domain (PF00248) - Aldo/keto reductase family	reductase	264909 264688, 18108362, 264558, 264600, 264760
1068	82082057 (2135, 2136)	Novel Protein sim. GBank gi 4599484 dbj BAA76770.1 - (AB023143) KIAA0928 protein [Homo sapiens]		UNCLASSIFIED	264604
1069	83002954 (2137, 2138)	Novel Protein sim. GBank gi 120304 sp P15932 FLGK_SALTY - FLAGELLAR HOOK-ASSOCIATED PROTEIN 1 (HAP1)		UNCLASSIFIED	264604, 264760
1070	82101992 (2139, 2140)	Novel Protein sim. GBank gi 1750397 U01261 - glutamate synthase large subunit [Pseudomonas aeruginosa]		synthase	264602
1071	20710589 (2141, 2142)	Novel Protein sim. GBank gi 477532 pir J A49175 - Molch B protein - mouse (fragment)	Contains protein domain (PF00008) - EGF-like domain	synthase	264606
1072	82356540 (2143, 2144)	Novel Protein sim. GBank gi 3893109 emb CAA76940 - (Y17820) CALO protein [Drosophila melanogaster]		UNCLASSIFIED	264687, 264688, 21908784, 35696052, 35695917, 35695855, 264600, 264601, 264602, 265009, 264605, 264508, 264905, 264690, 264906, 264782, 264628, 264788
1073	79814400 (2145, 2146)	Novel Protein sim. GBank gi 1176203 sp P46442 YHGM_ECOLI - HYPOTHETICAL 43.1 KD PROTEIN IN RPLM-HHOA INTERGENIC REGION (F375)		synthase	264909
1074	80105992 (2147, 2148)	Novel Protein sim. GBank gi 4033487 sp Q44472 TUD4_AGRVI - PUTATIVE HYDROXYPYRUVATE REDUCTASE		ATPase-associated	264769
1075	81850293 (2149, 2150)	Novel Protein sim. GBank gi 3413828 emb CAA20296 - (AL031260) hypothetical protein SC9A10.09 [Streptomyces coelicolor]		kinase	264905
1076	80477264 (2151, 2152)	Novel Protein sim. GBank gi 4033487 sp Q44472 TUD4_AGRVI - PUTATIVE HYDROXYPYRUVATE REDUCTASE		UNCLASSIFIED	264600
1077	79811334 (2153, 2154)	Novel Protein sim. GBank gi 4033487 sp Q44472 TUD4_AGRVI - PUTATIVE HYDROXYPYRUVATE REDUCTASE		UNCLASSIFIED	18108394, 264769, 264634, 264636
1078	20288874 (2155, 2156)	Novel Protein sim. GBank gi 4033487 sp Q44472 TUD4_AGRVI - PUTATIVE HYDROXYPYRUVATE REDUCTASE		UNCLASSIFIED	264600
1079	80494518 (2157, 2158)	Novel Protein sim. GBank gi 4033487 sp Q44472 TUD4_AGRVI - PUTATIVE HYDROXYPYRUVATE REDUCTASE		UNCLASSIFIED	18108394, 264769, 264634, 264636
1080	11767188 (2159, 2160)	Novel Protein sim. GBank gi 4033487 sp Q44472 TUD4_AGRVI - PUTATIVE HYDROXYPYRUVATE REDUCTASE		UNCLASSIFIED	264684
1081	94747080 (2161, 2162)	Novel Protein sim. GBank gi 4033487 sp Q44472 TUD4_AGRVI - PUTATIVE HYDROXYPYRUVATE REDUCTASE		UNCLASSIFIED	83373044, 265019, 22278002, 264482, 18108351, 264682, 264908, 264693, 264487, 264758, 264768, 264769, 21806767, 264511, 264910, 264634, 264635, 264905, 264636, 264906, 264837, 264907, 264908, 264764, 264638, 20281089, 264768, 264595
1082	81490656 (2163, 2164)	Novel Protein sim. GBank gi 4033487 sp Q44472 TUD4_AGRVI - PUTATIVE HYDROXYPYRUVATE REDUCTASE		UNCLASSIFIED	264684

1083	8745717 (2165, 2168)	Novel Protein sim. GBank gij1722945sp Q10523 YO1N_MYCTU - HYPOTHETICAL 44.8 KD PROTEIN CY427.23		UNCLASSIFIED	60424179, 264905, 264908, 264510, 60432229, 264759, 87168474, 264605, 264789, 264689, 18108384, 18108376, 35695855, 264638
1084	37799306 (2167, 2168)	Novel Protein sim. GBank gij418384sp P32057 WCAI_ECOLI - PUTATIVE COLANIC ACID BIOSYNTHESIS GLYCOSYL TRANSFERASE WCAI		UNCLASSIFIED	264769
1085	86475368 (2169, 2170)	Novel Protein sim. GBank gij1899180 (U90204) - heat shock protein 60 [Tukamurella tyrosinosolvens]	Contains protein domain (PF00118) - eph		60432229, 264687
1086	79608269 (2171, 2172)	Novel Protein sim. GBank gij1172858sp P46176 RL14_ACYKS - 50S RIBOSOMAL PROTEIN L14	Contains protein domain (PF00238) - ribosomalprot Ribosomal protein L14		264486
1087	79603978 (2173, 2174)	Novel Protein sim. GBank gij4160188 emb CAA15431 - (AL008583) dJ327.116.3 (novel CHROMObox family protein) [Homo sapiens]	Contains protein domain (PF00385) - chromo' (CHRomain Organization Modifier) domain		29331827, 264693
1088	79854963 (2175, 2176)	Novel Protein sim. GBank gij2983155 (AE000893) - phosphoglucomutase/phosphomannomutase [Aquifex aerophilus]		UNCLASSIFIED	264905, 264601, 18108387
1089	80216800 (2177, 2178)	Novel Protein sim. GBank gij4981768 gb AAD36290.1 AE001177 - (AE001178) NADH dehydrogenase, 30 kDa subunit, putative [Thermotoga maritima]	Contains protein domain (PF00329) - Respiratory-chain NADH dehydrogenase, 30 Kd subunit	UNCLASSIFIED	264488, 264511, 265011, 264682, 264768, 264689, 21908764, 35695917, 265020, 32833986, 18108370, 35695855
1090	11083825 (2179, 2180)	Novel Protein sim. GBank gij4007680 emb CAA223661 - (AL034443) putative oxidoreductase [Streptomyces coelicolor]			264604
1091	12917471 (2181, 2182)	Novel Protein sim. GBank gij2495582sp P7239 YLCD_ECOLI - HYPOTHETICAL 44.3 KD PROTEIN IN NFRB-PHEP INTERGENIC REGION PRECURSOR		UNCLASSIFIED	264637
1092	80252286 (2183, 2184)	Novel Protein sim. GBank gij2960098 emb CAA17966.11 - (AL022121) rth [Mycobacterium tuberculosis]	Contains protein domain (PF00730) - Endonuclease III	nuclease	264568 264769, 35695917, 35695855, 264600, 264602, 264603, 264605, 18108351
1093	80496304 (2185, 2186)	Novel Protein sim. GBank gij1001642 dbj BAA103731 - (D64002) dGTP triphosphohydrolase [Synecococcus sp.]		UNCLASSIFIED	264686
1094	10680972 (2187, 2188)	Novel Protein sim. GBank gij4585587 emb CAB40855.11 - (AL049828) putative adenine glycosylase [Streptomyces coelicolor]	Contains protein domain (PF00455) - Bacterial regulatory proteins, deoR family	nuclease	264908, 265007, 264595, 264600, 264602, 264603, 264604, 264605, 264762, 264768, 264769, 264636, 264558, 18108387,
1095	87457250 (2189, 2190)				60432113, 264482, 264486
1096	80025977 (2191, 2192)	Novel Protein sim. GBank gij115001sp P19206 BIOB_BACSH - BIOTIN SYNTHASE (BIOTIN SYNTHETASE)			264600, 264602, 264603, 264604, 264605, 264606, 264607, 35695917, 264682, 264631
1097	79239560 (2193, 2194)			UNCLASSIFIED	265019
1098	79186424 (2195, 2196)	Novel Protein sim. GBank gij114135sp P08205 ARGA_ECOLI - AMINO-ACID ACETYLTRANSFERASE (N-ACETYLGLUTAMATE SYNTHASE) (AGS)		synthase	264687

1099	39523838 (2197, 2198)	Novel Protein sim. GBank gi 3915144 sp O33017 TRMD_MYCLE - TRNA (GUANINE- N1)-METHYLTRANSFERASE (M1G- METHYLTRANSFERASE) (TRNA [GM37] METHYLTRANSFERASE)		UNCLASSIFIED	264603
1100	85736571 (2199, 2200)	Novel Protein sim. GBank gi 3023255 sp Q84420 ACOD_MESAU - ACYL-COA DESATURASE (STEAROYL-COA DESATURASE) (FATTY ACID DESATURASE) (DELTA(9)-DESATURASE)	desaturase		264259, 264636
1101	80491857 (2201, 2202)	Novel Protein sim. GBank gi 1174735 sp P43012 TOP1_HAEIN - DNA TOPOISOMERASE I (OMEGA-PROTEIN) (RELAXING ENZYME) (UNTWISTING ENZYME) (SWIVELASE)	Contains protein domain (PF01396) - Topoisomerase DNA binding C4 zinc finger	isomerase	264769
1102	79777614 (2203, 2204)	Novel Protein sim. GBank gi 1806596 sp U81788 - kinesin-73		UNCLASSIFIED	264910, 264909
1103	81897259 (2205, 2206)	[Drosophila melanogaster]		struct	264757
1104	95003115 (2207, 2208)	Novel Protein sim. GBank gi 2835448 sp AF048976 - synaptic ras GTPase-activating protein p135 SynGAP [Rattus norvegicus]		UNCLASSIFIED	26331822, 21908754, 264555, 264558, 264558, 22279002
1105	80255121 (2209, 2210)				264568
1106	79314110 (2211, 2212)			UNCLASSIFIED	264555, 264389
1107	80470019 (2213, 2214)				264908, 264769
1108	80440616 (2215, 2216)	Novel Protein sim. GBank gi 1173421 sp P43416 SECY_STRSC - PREPROTEIN TRANSLOCASE SECY SUBUNIT	transport		264807, 264510, 264511, 264600, 264602, 264605, 264768, 264769
1109	80084615 (2217, 2218)	Novel Protein sim. GBank gi 2895310 emb CAA18338 - (AL022268) putative ATP-dependent helicase (Streptomyces coelicolor)	helicase		264602, 264605, 264638
1110	80503554 (2219, 2220)				
1111	80071744 (2221, 2222)	Novel Protein sim. GBank gi 2622039 sp AE000668 - type I restriction modification system, subunit S [Methanobacterium thermoautotrophicum]			264908, 264593, 265010, 264601, 264603, 264604, 264605, 264682, 264769, 264893, 264836
1112	95010088 (2223, 2224)				18108370, 264557
1113	82456352 (2225, 2226)	Novel Protein sim. GBank gi 3216376 emb CAA18628 - (AL023862) putative oxidoreductase (Streptomyces coelicolor)	UNCLASSIFIED		264908 264600, 264602, 264604, 264605, 264762, 264769, 264585
1114	14998014 (2227, 2228)	Novel Protein sim. GBank gi 1083428 pir J554976 - NAD(P)+ transhydrogenase (B-specific) (EC 1.6.1.1) precursor - mouse	dehydrogenase		264636
1115	11765583 (2229, 2230)			UNCLASSIFIED	264686
1116	79841152 (2231, 2232)				264908

1117	95305465 (2233, 2234)	Novel Protein sim. GBank gij3255985[embjCAA94089] - (Z70200) US snRNP-specific 200KD protein [Homo sapiens]	Contains protein domain (PF00270) - helicase DEAD/DEAH box helicase	18108392, 264488, 263994, 264489, 18108398, 56182575, 22278995, 22278996, 35698288, 56994075, 22278997, 22278998, 22278999, 264259, 29331822, 56182181, 29331824, 66714117, 29331825, 29331826, 60432289, 29331827, 29331828, 35696052, 33656970, 28146498, 264508, 264905, 264909, 284906, 284907, 264908, 66712502, 264909, 52844045, 56182435, 264510, 264511, 265006, 265007, 264512, 265008, 265009, 264910, 60170831, 264591, 264592, 264593, 60433356, 264594, 60433438, 264595, 55812038, 264598, 264758, 264759, 21908754, 33109854, 33657084, 265011, 87168559, 264600, 264601, 264602, 264603, 265017, 264604, 265018, 264605, 265019, 264760, 55811150, 264881, 264782, 18108351, 264448, 264682, 264764, 264683, 264288, 264389, 264684, 264685, 264768, 264767, 264686, 264687, 264768, 264769, 264688, 18108359, 264689, 21906785, 21906788, 21906787, 21906788, 21908769, 55811957, 29148784, 35695917, 265020, 265021, 265022, 60170815, 264690, 264691, 33657023, 264692, 264693, 65274620, 33657109, 33657182, 27486281, 27486262, 27486284, 33657349, 27486285, 35695783, 264691
1118	79563328 (2235, 2236)		UNCLASSIFIED	264691
1119	79642463 (2237, 2238)		UNCLASSIFIED	264907
1120	79480463 (2239, 2240)	Novel Protein sim. GBank gij5420387[embjCAB46678.1] - (AJ243459) proteophosphoglycan [Leishmania major]	collagen	29331827, 265018, 265019, 264681, 265021, 60170615, 18108387
1121	79471716 (2241, 2242)	Novel Protein sim. GBank gij1644450 (U87884) - MEX-3 [Caenorhabditis elegans]	Contains protein domain (PF00013) - KH domain	264683, 264632, 18108388
1122	79456246 (2243, 2244)		UNCLASSIFIED	264639, 264563
1123	79637119 (2245, 2246)	Novel Protein sim. GBank gij98800[pirjS17768.3] - dehydroquinase synthase (EC 4.8.1.3) - Mycobacterium tuberculosis	synthase	264693, 27486285
1124	79811596 (2247, 2248)		UNCLASSIFIED	264909
1125	79757861 (2249, 2250)		UNCLASSIFIED	264910
1126	79758914 (2251, 2252)	Novel Protein sim. GBank gij138154[spjP03643]VGG_BPPHX - MAJOR SPIKE PROTEIN (G PROTEIN) (GPG)	eph	264905, 264909, 264910
1127	11800930 (2253, 2254)		UNCLASSIFIED	264682
1128	8384885 (2255, 2256)	Novel Protein sim. GBank gij5002704[embjCAB44358.1] - (AJ242630) DNA polymerase I [Methylobacterium sp. DM4]	Contains protein domain (PF00476) - DNA polymerase family A	264511

1129	80422480 (2257, 2258)	Novel Protein sim. GBank gij5689485dbj[BAA83028.1] - (AB028997) KIAA1074 protein [Homo sapiens]	Contains protein domain (PF00170) - bZIP transcription factor	UNCLASSIFIED	265011, 264768
1130	79420151 (2259, 2260)	Novel Protein sim. GBank gij4981328gbjAAD35881.1(AE00174 - (AE001747) bioY protein [Thermotoga maritima]		UNCLASSIFIED	264595
1131	80055391 (2281, 2282)	Novel Protein sim. GBank gij1841552 (U89336) - unknown (Homo sapiens)		UNCLASSIFIED	35696286, 22278998, 28331828, 264603, 264605, 264559
1132	82062248 (2263, 2264)	Novel Protein sim. GBank gij1841552 (U89336) - unknown (Homo sapiens)		UNCLASSIFIED	22278998, 264908, 265009, 264600, 264602, 264604, 264605, 264760, 32833986, 18108374
1133	17290437 (2285, 2286)	Novel Protein sim. GBank gij4982454gbjAAD36931.1(AE00182 - (AE001823) ATP-dependent protease LA, putative [Thermotoga maritima]		UNCLASSIFIED	265018
1134	80235376 (2267, 2268)	Novel Protein sim. GBank gij4539171embjCAB39700.1 - (AL049485) conserved hypothetical protein [Streptomyces coelicolor]		UNCLASSIFIED	264512, 264534
1135	80029393 (2289, 2270)	Novel Protein sim. GBank gij4539171embjCAB39700.1 - (AL049485) conserved hypothetical protein [Streptomyces coelicolor]		UNCLASSIFIED	264508, 264600, 264602, 264603, 18108376
1136	79842052 (2271, 2272)	Novel Protein sim. GBank gij4982454gbjAAD36931.1(AE00182 - (AE001823) ATP-dependent protease LA, putative [Thermotoga maritima]		UNCLASSIFIED	264908, 264908
1137	80931557 (2273, 2274)	Novel Protein sim. GBank gij4972746gbjAAD34788.1 - (AF132180) unknown [Drosophila melanogaster]	Contains protein domain (PF00515) - collagen TPR Domain		22278998, 22278999, 35696052, 264907, 265009, 60433356, 264598, 265010, 264448, 264882, 264767, 264688, 265020, 264692, 55811578, 35695855, 264631, 264632, 22278002, 264908
1138	79841163 (2275, 2276)	Novel Protein sim. GBank gij731807ispjP38739jYHC8_YEAST - HYPOTHETICAL 63.8 KD PROTEIN IN GUT1-RIM1 INTERGENIC REGION PRECURSOR	struc		
1139	79633561 (2277, 2278)	Novel Protein sim. GBank gij3650031 (AC005398) - putative proline-rich cell wall protein [Arabidopsis thaliana]		UNCLASSIFIED	264693
1140	39480358 (2279, 2280)	Novel Protein sim. GBank gij3650031 (AC005398) - putative proline-rich cell wall protein [Arabidopsis thaliana]		UNCLASSIFIED	264593
1141	79638019 (2281, 2282)	Novel Protein sim. GBank gij3650031 (AC005398) - putative proline-rich cell wall protein [Arabidopsis thaliana]		UNCLASSIFIED	265019, 264693
1142	19635848 (2283, 2284)	Novel Protein sim. GBank gij3650031 (AC005398) - putative proline-rich cell wall protein [Arabidopsis thaliana]		UNCLASSIFIED	264631
1143	87762158 (2285, 2286)	Novel Protein sim. GBank gij3628000jembjCAA05880 - (AJ003125) procollagen I N-proteinase [Homo sapiens]	Contains protein domain (PF00090) - Thrombospondin type 1 domain		56182575, 264908, 264600, 264632, 87168518
1144	80068988 (2287, 2288)	Novel Protein sim. GBank gij3628000jembjCAA05880 - (AJ003125) procollagen I N-proteinase [Homo sapiens]		UNCLASSIFIED	264635, 264638, 264907, 264593, 264908, 264568, 264909
1145	14610262 (2289, 2290)	Novel Protein sim. GBank gij3628000jembjCAA05880 - (AJ003125) procollagen I N-proteinase [Homo sapiens]		UNCLASSIFIED	264112
1146	82062082 (2281, 2292)	Novel Protein sim. GBank gij3628000jembjCAA05880 - (AJ003125) procollagen I N-proteinase [Homo sapiens]		UNCLASSIFIED	264769, 264689, 35698286, 264760, 264905, 264486, 264559
1147	80071761 (2293, 2294)	Novel Protein sim. GBank gij2499003ispjP76422jTHID_ECOLI - PHOSPHOMETHYL PYRIMIDINE KINASE (HMP-PHOSPHATE KINASE) (HMP-P KINASE)	kinase		264557
1148	80048433 (2295, 2296)	Novel Protein sim. GBank gij2499003ispjP76422jTHID_ECOLI - PHOSPHOMETHYL PYRIMIDINE KINASE (HMP-PHOSPHATE KINASE) (HMP-P KINASE)			264591
1149	11607438 (2297, 2298)	Novel Protein sim. GBank gij2698734jembjCAA17213.1 - (AL021897) hypothetical protein Rv1097c [Mycobacterium tuberculosis]			264591

1150	81325074 (2289, 2300)	Novel Protein sim. GBank gll2895095 (AF011337) - putative E1-E2 ATPase [Mus musculus]		ATPase-associated	264488, 35698288, 284907, 264908, 264909, 264910, 284593, 284596, 284758, 264764, 284768, 284768, 284693, 284628, 60431850, 284584, 284568, 284587
1151	80070874 (2301, 2302)	Novel Protein sim. GBank gll4324655(gb AAD16978) - (AF108191) DNA polymerase III alpha subunit [Streptomyces coelicolor]		polymerase	264595
1152	80235547 (2303, 2304)	Novel Protein sim. GBank gll3874275(emb CA807311.1) - (Z82825) predicted using GeneFinder; Similarity to Yeast low affinity glucose transporter HXT4 (PS:32467); cDNA EST EMBL.C12555 comes from this gene; cDNA EST yk404c10.3 comes from this gene; cDNA EST yk404c10.5 comes from thi...		glycoprotein	284488, 22278988, 284905, 284629, 264486
1153	80027783 (2305, 2308)	Novel Protein sim. GBank gll4240315(db BAA74938.1) - (AB020720) KIAA0913 protein [Homo sapiens]		UNCLASSIFIED	264910, 284555, 284557
1154	83002895 (2307, 2308)				265008
1155	78411098 (2309, 2310)	Novel Protein sim. GBank gll586855(sp P37617 ATZN_ECOLI - ZINC-TRANSPORTING ATPASE (ZINC)-TRANSLLOCATING P-TYPE ATPASE)	Contains protein domain (PF00122) - transport E1-E2 ATPase	UNCLASSIFIED	264690, 284638
1156	57147843 (2311, 2312)				264603
1157	95287711 (2313, 2314)	Novel Protein sim. GBank gll418480(sp P32139 VHR_ECOLI - HYPOTHETICAL 34.0 KD PROTEIN IN GLNA-RBN INTERGENIC REGION)		UNCLASSIFIED	284906, 264907, 264758, 284768, 264769, 284689, 284638, 284588
1158	82454817 (2315, 2316)	Novel Protein sim. GBank gll2483481(sp Q50724 Y09S_MYCTU - HYPOTHETICAL 87.3 KD PROTEIN CY78.27C)			264908, 284762, 284687, 264769, 284689, 18108374, 35698555
1159	79188451 (2317, 2318)	Novel Protein sim. GBank gll1138408(db BAA11490) - (D78995) similar to pig tubulin-tyrosine ligase. [Homo sapiens]			264687
1160	91228893 (2319, 2320)				29331822, 29331824, 68714117, 29331828, 60433358, 265018, 265019, 83373044, 18108385, 22279000, 22279002, 264563
1161	7417143 (2321, 2322)	Novel Protein sim. GBank gll2443342(db BAA22380) - (D88764) alpha 2 type I collagen [Rana catesbeiana]		UNCLASSIFIED	264389
1162	78635357 (2323, 2324)	Novel Protein sim. GBank gll4503375(ref NP_001376.1 pDPYS - dihydropyrimidinase)		UNCLASSIFIED	264693
1163	78563186 (2325, 2326)	Novel Protein sim. GBank gll5052554(gb AAD38607.1 AF145632) - (AF145632) BcDNA, GH06032 [Drosophila melanogaster]		transport	29331827, 264908
1164	78650829 (2327, 2328)				18108388, 29331827, 29331828, 29146498, 29146499, 18108354, 21906788, 29148827, 21908789, 284693, 18108382, 18108385
1165	80491888 (2329, 2330)			UNCLASSIFIED	264802, 284605, 264769, 18108370, 18108374, 284565
1166	88098456 (2331, 2332)	Novel Protein sim. GBank gll4589476(db BAA78768.1) - (AB023138) KIAA0922 protein [Homo sapiens]			264488, 35698288, 22278989, 284259, 68714117, 60432289, 35698052, 284905, 56192435, 265008, 60433438, 284759, 21906754, 33109954, 265017, 265019, 284448, 284288, 284768, 284685, 35698423, 35698555, 264558, 18108385, 60432113

1167	79863862 (2333, 2334)	Novel Protein sim. GBank gij2580433[dj]BAA23138] - (D76414) ppGpp hydrolase [Staphylococcus aureus]		Kinase	264488
1168	88094678 (2335, 2336)			UNCLASSIFIED	264259, 29331827, 56182435, 60433438, 285019, 33857023, 35698555, 264568
1169	11805403 (2337, 2338)			UNCLASSIFIED	264681
1170	21632244 (2339, 2340)			UNCLASSIFIED	264802
1171	20434582 (2341, 2342)	Novel Protein sim. GBank gij2727814 (AF029249) - procollagen D [Mytilus edulis]		UNCLASSIFIED	264556
1172	78610113 (2343, 2344)	Novel Protein sim. GBank gij4757846[ref]NP_004317.1p8CL9 - B-cell CLL/lymphoma 9		UNCLASSIFIED	55810764, 35696052, 264634, 264486
1173	80235713 (2345, 2346)	Novel Protein sim. GBank gij2584053[dj]BAA22846] - (AB007832) Bm trachealless [Bombyx mori]			264508, 284906, 264907, 264809, 264591, 264832, 264638, 264839
1174	20293077 (2347, 2348)	Novel Protein sim. GBank gij2911027[emb]CAA17520] - (AL021958) mmsA [Mycobacterium tuberculosis]		dehydrogenase	264600
1175	20711847 (2348, 2350)	Novel Protein sim. GBank gij118333[sp]P23234[DCIP ENTCL - INDOLE-3-PYRUVATE DECARBOXYLASE (INDOLEPYRUVATE DECARBOXYLASE)]	Contains protein domain (PF00205) - Thiamine pyrophosphate enzymes	carboxylase	264601
1176	80252845 (2351, 2352)	Novel Protein sim. GBank gij1144520 (U34956) - phosphoribosylformylglycinamide synthase [Mycobacterium tuberculosis]	Contains protein domain (PF00586) - AIR synthase related protein	synthase	264509, 264805, 264593, 264802, 264605
1177	80064647 (2353, 2354)	Novel Protein sim. GBank gij118791[sp]P28643[FABG, CUPLA - 3-OXOACYL-ACYL-CARRIER PROTEIN] REDUCTASE PRECURSOR (3-KETOACYL-ACYL CARRIER PROTEIN REDUCTASE)	Contains protein domain (PF00106) - short chain dehydrogenase	reductase	264605
1178	94128641 (2355, 2356)	Novel Protein sim. GBank gij5031697[ref]NP_005594.1pFIC1 - familial intrahepatic cholestasis 1, (progressive, Byler disease and benign recurrent)	Contains protein domain (PF00122) - E1-E2 ATPase		65274572, 18108398, 22278998, 22278999, 29331828, 264508, 264908, 264828, 33857402, 33108954, 264769, 21808765, 21808786, 21808788, 55811957, 33857023, 264629, 55811576, 35698423, 264636, 264556, 56182323, 60432113, 22279000, 22279002
1179	80055575 (2357, 2358)	Novel Protein sim. GBank gij2980090[emb]CAA17988.1] - (AL022121) dppA [Mycobacterium tuberculosis]	Contains protein domain (PF00496) - Bacterial extracellular solute-binding proteins, family 5	transport	264603
1180	11784446 (2359, 2360)	Novel Protein sim. GBank gij2558614[emb]CAA04787] - (AJ001493) dehydroquinase dehydratase [Streptomyces coelicolor]	Contains protein domain (PF01220) - Dehydroquinase class II	synthase	264638
1181	17846382 (2381, 2382)			UNCLASSIFIED	265017
1182	81494284 (2383, 2384)	Novel Protein sim. GBank gij5420387[emb]CAB46678.1] - (AJ243458) proteophosphoglycan [Leishmania major]			265007, 265009, 264564, 264908, 264693
1183	78574044 (2365, 2366)				
1184	52558933 (2367, 2368)	Novel Protein sim. GBank gij4091877 (AF061331) - alpha galactosidase precursor [Saccharopolyspora erythraea]		UNCLASSIFIED	264689, 35698423, 264638, 18108385
1185	78491185 (2369, 2370)	Novel Protein sim. GBank gij2128478[sp]IS1939 - chitinase (EC 3.2.1.14) precursor - bee!		UNCLASSIFIED	264802
				glycoprotein	263987

1188	20224012 (2371, 2372)				UNCLASSIFIED	284559
1187	79248834 (2373, 2374)				UNCLASSIFIED	2931825, 285017, 18108351
1186	78831387 (2375, 2376)	Novel Protein sim. GBank gij2896039 (AF054525) - hypothetical protein [Synectococcus PCC7002]			UNCLASSIFIED	284905, 284908
1185	78609367 (2377, 2378)					284692
1184	78930589 (2379, 2380)				UNCLASSIFIED	285018
1183	80310105 (2381, 2382)				UNCLASSIFIED	284600, 284605, 284764, 35695855, 284638, 284488
1182	13521641 (2383, 2384)		Contains protein domain (PF01561) - FMRFamide related peptide family			284636
1181	11103584 (2385, 2386)				UNCLASSIFIED	283978
1180	78830947 (2387, 2388)	Novel Protein sim. GBank gij854085jemb[CAA58337] - (X83413) U88 [Human herpesvirus 8]			UNCLASSIFIED	285007, 285008
1179	20445442 (2389, 2390)	Novel Protein sim. GBank gij1780277 (AE000459) - putative oxidoreductase [Escherichia coli]			UNCLASSIFIED	284605
1178	13000668 (2391, 2392)					284689
1177	11392317 (2393, 2394)	Novel Protein sim. GBank gij2497360[spQ50715]JMDH_MYCTU - INOSINE-5'- MONOPHOSPHATE DEHYDROGENASE (IMP DEHYDROGENASE) (IMPDH) (IMPD)	Contains protein domain (PF00571) - CBS domain		dehydrogenase	284594
1176	95280101 (2395, 2396)					284603
1175	81882011 (2397, 2398)	Novel Protein sim. GBank gij1709525[spP54673]P3K1_DICD1 - PHOSPHATIDYLINOSITOL 3-KINASE 1 (PI3-KINASE) (PTDINS-3-KINASE) (PI3K)				284259, 284757, 33109854, 21908788
1200	9848880 (2399, 2400)					
1201	80503751 (2401, 2402)	Novel Protein sim. GBank gij2499877[spP70845]BLMH_RAT - BLEOMYCIN HYDROLASE (BLM HYDROLASE) (BMH)			UNCLASSIFIED cathepsin	284910 284766, 284769
1202	80082633 (2403, 2404)	Novel Protein sim. GBank gij608342 (U18997) - ORF_0622; reading frame open far upstream of start; possible frameshift, linking to previous ORF [Escherichia coli]			ribosomal prot	284600, 284558
1203	82125373 (2405, 2406)				UNCLASSIFIED	284768, 284769, 35695917, 284910, 284780, 284908, 284907, 284629, 284908, 284909, 284768
1204	80503916 (2407, 2408)	Novel Protein sim. GBank gij2500728[spQ59912]SECY_STRGB - PREPROTEIN TRANSLOCASE SECY SUBUNIT				284905, 284769, 284636
1205	80053961 (2409, 2410)				UNCLASSIFIED	284568
1206	80241985 (2411, 2412)				UNCLASSIFIED	284556, 284557, 284558
1207	78841192 (2413, 2414)					2931824, 284909, 285021, 18108370
1208	87755217 (2415, 2416)	Novel Protein sim. GBank gij2845560 (AF027954) - Bcl-2- related ovarian killer protein [Rattus norvegicus]	Contains protein domain (PF00452) - Apoptosis regulator proteins, Bcl-2 family		apoptosis	2931824, 29331825, 29331827, 285007, 284764, 284683, 284769, 284688, 284689

1209	78185742 (2417, 2418)	Novel Protein sim. GBank gij1175033 p P44398 XYLA_HAEN - XYLOSE ISOMERASE	Contains protein domain (PF00259) - Xylose isomerase	Isomerase	264887, 264888
1210	56428884 (2419, 2420)				
1211	94865855 (2421, 2422)	Novel Protein sim. GBank gij421095 p I S30888 - hypothetical protein o248 - Escherichia coli		UNCLASSIFIED transferase	264907, 264993 284591, 284592, 284595
1212	78187929 (2423, 2424)	Novel Protein sim. GBank gij3880625 emb CAB07858 - (Z33765) predicted using GeneFinder; similar to RNA recognition motif. (aka RRM, RBD, or RNP domain); cDNA EST EMBL: T01682 comes from this gene; cDNA EST EMBL: M75823 comes from this gene; cDNA EST EMBL: D27559 comes from this ge...	Contains protein domain (PF01412) - Putative GTP-ase activating protein for Arf		264689, 263967
1213	78858633 (2425, 2426)	Novel Protein sim. GBank gij228292 p I J1503375A - vir gene [Bordetella pertussis]		Kinase	264909
1214	10144308 (2427, 2428)	Novel Protein sim. GBank gij5726285 gb AAD48396.1 AF12616 - (AF126162) HERV-H LTR associating protein 2 [Homo sapiens]		UNCLASSIFIED	264908
1215	80050106 (2429, 2430)	Novel Protein sim. GBank gij2326739 emb CAB10953 - (Z98268) recN [Mycobacterium tuberculosis]		UNCLASSIFIED	265008, 264601, 264602, 264603, 33657109
1216	20438324 (2431, 2432)	Novel Protein sim. GBank gij417329 p P33038 MURA_ENTCL - UDP-N- ACETYLGLUCOSAMINE 1- CARBOXYVINYLTRANSFERASE (ENOYLPIRUVATE TRANSFERASE) (UDP-N-ACETYLGLUCOSAMINE ENOLPYRUVYL TRANSFERASE) (EPT)		transferase	264604
1217	95011344 (2433, 2434)				
1218	11093880 (2435, 2436)	Novel Protein sim. GBank gij1805460 dbj BAA09022 - (D50453) homologue of succinate semialdehyde dehydrogenase GabD of E. coli [Bacillus subtilis]		UNCLASSIFIED	264905, 264907, 264908, 264591, 264768, 264691, 264693, 264628, 264630, 264636, 264564
1219	81216252 (2437, 2438)			dehydrogenase	264601
1220	81241524 (2439, 2440)	Novel Protein sim. GBank gij4240315 dbj BAA74938.1 - (AB020720) KIAA0913 protein [Homo sapiens]		oncogene	56181688, 29331822, 60432289, 264601, 264692, 264629
1221	83045055 (2441, 2442)	Novel Protein sim. GBank gij2143886 p I J52523 - nucleoporin p82 homolog - rat (fragment)		UNCLASSIFIED	52844507, 264805, 264909, 265008, 265019, 265020, 52844150, 33657023, 264693, 33857182, 35895763, 264634, 22279000, 22279002, 264482
1222	20711865 (2443, 2444)	Novel Protein sim. GBank gij730805 p P39663 SPHR_SYN7 - ALKALINE PHOSPHATASE SYNTHESIS TRANSCRIPTIONAL REGULATORY PROTEIN SPHR	Contains protein domain (PF00486) - Transcriptional regulatory protein, C terminal	phosphatase	264768, 265020, 264906
1223	11615647 (2445, 2446)				264601
1224	80432845 (2447, 2448)	Novel Protein sim. GBank gij1172627 p P46546 PROB_CORGL - GLUTAMATE 5- KINASE (GAMMA-GLUTAMYL KINASE) (GK)	Contains protein domain (PF01472) - PUA domain	kinase	264593 264593, 264600, 264601, 284603, 264605, 264768, 18108376, 264635, 18108387

1225	80434427 (2449, 2450)	Novel Protein sim. GBank gij2105050[embjCAB08836] - (Z95436) hypothetical protein RV3644c [Mycobacterium tuberculosis]			264768
1226	80237518 (2451, 2452)	Novel Protein sim. GBank gij1706768[spP98133]FBN1_BOVIN - FIBRILLIN 1 PRECURSOR (MP340)		polymerase	264905, 264512, 264689
1227	79422138 (2453, 2454)	Novel Protein sim. GBank gij1653901[dbjBAA18811] - (D90917) acriflavine resistance protein [Synecchocystis sp.]		UNCLASSIFIED	264908, 264637, 264639
1228	79209027 (2455, 2456)	Novel Protein sim. GBank gij1653901[dbjBAA18811] - (D90917) acriflavine resistance protein [Synecchocystis sp.]	Contains protein domain (PF00873) - AcrB/AcrD/AcrF family		264605, 264634
1229	94329135 (2457, 2458)	Novel Protein sim. GBank gij116230[spP28598]CH60_BACSU - 60 KD CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN)	Contains protein domain (PF00118) - TCP-1/cpn60 chaperonin family	UNCLASSIFIED	87169474, 265011, 87168559, 264681, 264689, 264693, 85274820, 18108374
1230	80049357 (2459, 2460)	Novel Protein sim. GBank gij116230[spP28598]CH60_BACSU - 60 KD CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN)		eph	264908, 264605, 18108388
1231	79843141 (2461, 2462)	Novel Protein sim. GBank gij1215733 (U48718) - OphC [Agrobacterium tumefaciens]		UNCLASSIFIED	264908
1232	79853104 (2463, 2464)	Novel Protein sim. GBank gij1215733 (U48718) - OphC [Agrobacterium tumefaciens]		transport	264909
1233	80255179 (2465, 2466)	Novel Protein sim. GBank gij116288[spP20730]CHHC_BOMMO - CHORION CLASS HIGH-CYSTEINE HCB PROTEIN 13 PRECURSOR (HC-B.13)		UNCLASSIFIED	265017, 264564
1234	79242158 (2467, 2468)	Novel Protein sim. GBank gij729871[spP40280]HZA_MAIZE - HISTONE H2A	Contains protein domain (PF00125) - Core histone H2A/H2B/H3/H4	histone	265008, 265010, 18108381
1235	79914423 (2469, 2470)	Novel Protein sim. GBank gij3875133[embjCAA94750] - (Z70750) similar to actin binding domain; cDNA EST EMBL:T00093 comes from this gene; cDNA EST EMBL:D34443 comes from this gene; cDNA EST EMBL:D37508 comes from this gene; cDNA EST EMBL:D64247 comes from this gene; cDNA EST EMBL...		UNCLASSIFIED	264634, 264762
1236	81927147 (2471, 2472)	Novel Protein sim. GBank gij3875133[embjCAA94750] - (Z70750) similar to actin binding domain; cDNA EST EMBL:T00093 comes from this gene; cDNA EST EMBL:D34443 comes from this gene; cDNA EST EMBL:D37508 comes from this gene; cDNA EST EMBL:D64247 comes from this gene; cDNA EST EMBL...			265018, 55811150, 264565, 264757
1237	83371782 (2473, 2474)	Novel Protein sim. GBank gij3875133[embjCAA94750] - (Z70750) similar to actin binding domain; cDNA EST EMBL:T00093 comes from this gene; cDNA EST EMBL:D34443 comes from this gene; cDNA EST EMBL:D37508 comes from this gene; cDNA EST EMBL:D64247 comes from this gene; cDNA EST EMBL...			264758, 264601, 264766, 264687, 18108372, 264555, 264559
1238	87411577 (2475, 2476)	Novel Protein sim. GBank gij3885470 (AF061443) - G protein-coupled receptor LGR4 [Rattus norvegicus]	Contains protein domain (PF00560) - Leucine Rich Repeat	glycoprotein	264259, 29331822, 29331824, 35686052, 264508, 284908, 52844045, 52846317, 264288, 264769, 264693, 264632, 264634, 264558, 87168518, 264563
1239	82197449 (2477, 2478)	Novel Protein sim. GBank gij4007890[gbjAAC95339] - (AF084383) DOK protein [Mus musculus]		oncogene	264509, 264511, 264759, 264760, 264764, 264557
1240	80497259 (2479, 2480)	Novel Protein sim. GBank gij1176192[spP45420]YHCD_ECOLI - HYPOTHETICAL OUTER MEMBRANE USHER PROTEIN IN GLTF-NANT INTERGENIC REGION PRECURSOR			264769
1241	80020711 (2481, 2482)	Novel Protein sim. GBank gij121383[spP19904]GLNA_VIBAL - GLUTAMINE SYNTHETASE (GLUTAMATE-AMMONIA LIGASE)	Contains protein domain (PF00120) - Glutamine synthetase	UNCLASSIFIED	264601, 264604, 264638
1242	79775690 (2483, 2484)	Novel Protein sim. GBank gij121383[spP19904]GLNA_VIBAL - GLUTAMINE SYNTHETASE (GLUTAMATE-AMMONIA LIGASE)			264908, 264907, 264908, 264634

1243	78779458 (2485, 2486)	Novel Protein sim. GBank glj335567 [emb]CAA18971] - (AL031124) branched-chain amino acid aminotransferase [Streptomyces coelicolor]		UNCLASSIFIED	18108374, 35695917, 35695855, 265009, 264508, 264809
1244	10284821 (2487, 2488)	Novel Protein sim. GBank glj2870848 (AF051845) - Xin [Mus musculus]		UNCLASSIFIED	264691
1245	80437103 (2489, 2490)	Novel Protein sim. GBank glj4588330 [gb]BAA76357.1] - (AB016787) cytochrome o ubiquinol oxidase B [Pseudomonas putida]	Contains protein domain (PF00115) Cytochrome C and Quinol oxidase polypeptide I	oxidase	264768
1246	80059321 (2491, 2492)	Novel Protein sim. GBank glj3581849 [emb]CAA20805] - (AL031541) putative phenylalanyl-tRNA synthetase beta chain [Streptomyces coelicolor]		UNCLASSIFIED	264604, 264836, 264557, 264584
1247	80064831 (2493, 2494)	Novel Protein sim. GBank glj2821684 (AE000842) - adhesion protein [Methanobacterium thermoautotrophicum]			264758, 264605, 264639
1248	88070353 (2495, 2496)	Novel Protein sim. GBank glj1352403 [sp]P09467 [F16P_HUMAN - FRUCTOSE-1,6-BISPHOSPHATASE (D-FRUCTOSE-1,6-BISPHOSPHATE 1-PHOSPHOHYDROLASE) (FBPASE)]	Contains protein domain (PF00316) - Fructose-1,6-bisphosphatase	UNCLASSIFIED	18108392, 264259, 29331826, 264108, 264508, 264807, 264828, 265009, 80433356, 264757, 264758, 21808754, 265010, 265011, 265018, 265019, 264760, 18108351, 18108354, 265021, 18108376, 18108377, 264630, 18108385
1249	80056657 (2497, 2498)	Novel Protein sim. GBank glj2791407 [emb]CAA18001] - (AL021184) hypothetical protein Rv1473 [Mycobacterium tuberculosis]		transport	264908, 265010, 264600, 264603, 264691, 18108378
1250	12694385 (2498, 2500)	Novel Protein sim. GBank glj112785 [sp]P05100 [3MG1 - DNA-3-METHYLADENINE GLYCOSIDASE I (3-METHYLADENINE-DNA GLYCOSYLASE I, CONSTITUTIVE) (TAG I)]		UNCLASSIFIED	264689
1251	78850448 (2501, 2502)				264809
1252	79458087 (2503, 2504)			UNCLASSIFIED	264683, 263978
1253	80050121 (2505, 2506)	Novel Protein sim. GBank glj5870178 [gb]AAD46816.1 [AF16131 - (AF161317) NRAMP manganese transport protein MnTA [Salmonella typhimurium]		glycoprotein	264800, 264803, 18108376
1254	87716767 (2507, 2508)	Novel Protein sim. GBank glj103180 [pir]S22128 - finger protein unkempt - fruit fly [Drosophila melanogaster]		UNCLASSIFIED	35696286, 264910, 264764, 264688, 21906787, 55811957, 264892, 264556, 264639
1255	79169728 (2509, 2510)				264638
1256	87889508 (2511, 2512)	Novel Protein sim. GBank glj2985353 [emb]CAA04608.1] - (AJ001206) pep2 [Streptomyces coelicolor]		UNCLASSIFIED	60432289, 264600, 264605, 264764, 264687, 264769, 264889, 27488265, 18108374, 18108378
1257	80201435 (2513, 2514)	Novel Protein sim. GBank glj3183306 (AF069300) - contains similarity to Arabidopsis membrane-associated salt inducible-like protein (GB:AL021637) [Arabidopsis thaliana]		UNCLASSIFIED	264094, 265019
1258	20708150 (2515, 2516)				264602, 263978
1259	80186012 (2517, 2518)			UNCLASSIFIED	264806, 264448, 264908
1260	80084608 (2519, 2520)			UNCLASSIFIED	264634, 264639

1261	87412802 (2521, 2522)	Novel Protein sim. GBank gl 568951 dbj BAA83039.1 - (AB029010) KIAA1087 protein [Homo sapiens]	Contains protein domain (PF01699) - cadherin	29331824, 264906, 264909, 264768, 264769, 264689, 264693, 264939, 18108384, 284583
1262	13504589 (2523, 2524)	Novel Protein sim. GBank gl 95100 pir IS21334 - hypothetical protein 4 - Agrobacterium tumefaciens	UNCLASSIFIED	264634
1263	20710997 (2525, 2526)	Novel Protein sim. GBank gl 3550958 (AF004840) - CDO [Rattus norvegicus]	struct	264602
1264	80083386 (2527, 2528)			264634
1265	80253578 (2529, 2530)		UNCLASSIFIED	264563
1266	78914604 (2531, 2532)		UNCLASSIFIED	264768, 264638, 264638, 264567
1267	80538918 (2533, 2534)	Novel Protein sim. GBank gl 1085002 pir IS55056 - mitochondrial carrier protein DIF-1 homolog - Caenorhabditis elegans	transport	264259, 21906754, 264369
1268	88178473 (2535, 2536)	Novel Protein sim. GBank gl 4888445 emb CAB43370.1 - (AL050269) hypothetical protein [Homo sapiens]	Contains protein domain (PF00583) - Acetyltransferase (GNAT) family	18108398, 22278995, 56994075, 60424289, 29331827, 264108, 264512, 265007, 265008, 265009, 264595, 33109954, 33657084, 87188559, 264600, 265018, 265019, 264369, 264888, 21906787, 265020, 52844150, 264891, 33657023, 33657349, 18108374, 264556, 18108385, 80432113, 22278002, 264488
1269	78821848 (2537, 2538)	Novel Protein sim. GBank gl 3334791 emb CAA18939 - (AL031107) hypothetical protein SC5A7.10c [Streptomyces coelicolor]	UNCLASSIFIED	264508, 264905, 264906, 264687, 264693
1270	80031420 (2539, 2540)	Novel Protein sim. GBank gl 2851634 sp Q50591 Y0D1_MYCTU - HYPOTHETICAL 50.0 KD PROTEIN CY1A11.01	Contains protein domain (PF01574) - IMP dehydrogenase / GMP reductase N terminus	265010, 264601
1271	78840499 (2541, 2542)		ATPase associated	35886052, 264908
1272	79482878 (2543, 2544)			264688, 264689
1273	60220315 (2545, 2546)	Novel Protein sim. GBank gl 1855685 emb CAB03731 - (Z81368) hypothetical protein Rv2395 [Mycobacterium tuberculosis]	UNCLASSIFIED	264509, 264639
1274	95010802 (2547, 2548)		UNCLASSIFIED	
1275	20730763 (2549, 2550)	Novel Protein sim. GBank gl 123728 sp P10413 HTPG_ECOLI - HEAT SHOCK PROTEIN HTPG (HIGH TEMPERATURE PROTEIN G) (HEAT SHOCK PROTEIN C82.5)	Contains protein domain (PF00183) - Hsp90 protein	264905, 264908, 264909, 264769
1276	21148644 (2551, 2552)	Novel Protein sim. GBank gl 2129478 pir IS51939 - chitinase (EC 3.2.1.14) precursor - beet	UNCLASSIFIED	264602
1277	20438195 (2553, 2554)			
1278	11088365 (2555, 2556)	Novel Protein sim. GBank gl 175473 sp P44555 YAAJ_HAEIN - HYPOTHETICAL PROTEIN HI0183	UNCLASSIFIED	264558
1279	21658756 (2557, 2558)	Novel Protein sim. GBank gl 1929513 (U64318) - ATP synthase subunit beta [Moodella thermacetica]	UNCLASSIFIED	264603
1280	78310959 (2559, 2560)	Novel Protein sim. GBank gl 4938504 emb CAB43882.1 - (AL078465) putative protein [Arabidopsis thaliana]	synthase	264605
			struct	263878

1281	94323988 (2561, 2562)	Novel Protein sim. GBank gll1136501 (U39546) - surface protein MCA-32 [Rattus norvegicus]	Contains protein domain (PF00047) - immunoglobulin domain	UNCLASSIFIED	29331825, 29331828, 264766, 83373044
1282	87537695 (2563, 2564)	Novel Protein sim. GBank gll3328180 (AF074266) - proto-oncogene AF-4 [Mus musculus]		UNCLASSIFIED	285008
1283	20466305 (2565, 2566)	Novel Protein sim. GBank gll3281721 (emb) CAB070571 - (Z92770) hypothetical protein Rv0153c [Mycobacterium tuberculosis]		UNCLASSIFIED	284605
1284	20636325 (2567, 2568)	Novel Protein sim. GBank gll3929022 (AF057696) - LspB [Haemophilus ducreyi]			284604
1285	80427330 (2569, 2570)	Novel Protein sim. GBank gll417154 (sp) P33126 (HS2_ORYSA - HEAT SHOCK PROTEIN 82)	Contains protein domain (PF00183) - eph Hsp90 protein		284766, 264689, 263967
1286	20465254 (2571, 2572)	Novel Protein sim. GBank gll2078004 (emb) CAB084511 - (Z95237) gorA [Mycobacterium tuberculosis]	reductase		264605, 264639
1287	80417530 (2573, 2574)			UNCLASSIFIED	265011, 284602, 264766, 264687, 264769, 264689, 18108370, 264638, 18108385, 264563
1288	95338101 (2575, 2576)	Novel Protein sim. GBank gll5353510 (gb) AAD42161.1 (AF08891 - (AF088916) emilin precursor [Homo sapiens])	Contains protein domain (PF00386) - C1q domain	collagen	35896032, 264107, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 285007, 264512, 264910, 265009, 33857402, 264595, 264758, 265011, 265019, 264760, 18108351, 264681, 264784, 264288, 264885, 264768, 264687, 264768, 264769, 265020, 285021, 264534, 264692, 18108370, 264628, 18108374, 35896423, 264555, 264556, 264557, 264558, 18108385, 264564, 264566, 264567, 264688, 18108391
1289	11813647 (2577, 2578)	Novel Protein sim. GBank gll1169995 (sp) P46023 (GPCR_LYMST - G-PROTEIN COUPLED RECEPTOR GRL101 PRECURSOR)	Contains protein domain (PF00001) - 7 transmembrane receptor (rhodopsin family)	UNCLASSIFIED	284637
1290	19526027 (2579, 2580)	Novel Protein sim. GBank gll2072674 (emb) CAB083051 - (Z95120) rhlE [Mycobacterium tuberculosis]	Contains protein domain (PF00271) - Helicases conserved C-terminal domain	lm7	284563
1291	80470266 (2581, 2582)	Novel Protein sim. GBank gll1635755 (U86338) - zinc finger protein Png-1 [Mus musculus]	Helicases conserved C-terminal domain	helicase	265007, 265008, 264769
1292	94723316 (2583, 2584)	Novel Protein sim. GBank gll2129173 (pir) JF8453 - oxaloacetate decarboxylase (EC 4.1.1.3) alpha subunit - Melhanococcus jannaschii	Contains protein domain (PF01530) - Zinc finger, C2HC type	transcript factor	264092, 264259, 29331822, 29331824, 264508, 264906, 264909, 264512, 265008, 265009, 264591, 265019, 264389, 264288, 264686, 264768, 264693, 18108374, 264632, 58182323, 264639, 83373044, 22279002, 264482, 264563
1293	80067536 (2585, 2586)	Novel Protein sim. GBank gll2129173 (pir) JF8453 - oxaloacetate decarboxylase (EC 4.1.1.3) alpha subunit - Melhanococcus jannaschii		UNCLASSIFIED	265008, 55812038, 264389, 264556
1294	82125908 (2587, 2588)	Novel Protein sim. GBank gll5441779 (emb) CAB46803.11 - (AL096811) putative alcohol dehydrogenase (zinc-binding) [Streptomyces coelicolor A3(2)]	biolindap		264602, 264605, 264760, 18108351, 264689, 33857023, 264559
1295	11686851 (2589, 2590)		dehydrogenase		264689

1296	11887804 (2581, 2592)	Novel Protein sim. GBank	UNCLASSIFIED	284591, 284639	
1297	78639300 (2593, 2594)	Novel Protein sim. GBank gl4982191gb AAD36886.1 AE001805 DNA- directed DNA polymerase I [Thermotoga maritima]	Contains protein domain (PF01367) - 5'-3' exonuclease	284693	
1298	94239508 (2595, 2596)	Novel Protein sim. GBank gl1943770 (U97191) - F53F10.1 gene product [Caenorhabditis elegans]	struct	18108348, 265017	
1299	80255378 (2597, 2598)	Novel Protein sim. GBank gl3445181 (AC005498) - R31685.2 [Homo sapiens]		284488, 284906, 284909, 22279002, 284586	
1300	80064867 (2599, 2600)	Novel Protein sim. GBank gl4082973 gb BAA36204.1 - (AB017138) alpha subunit of malonate decarboxylase [Pseudomonas putida]	transcriptfactor	284605	
1301	17839614 (2601, 2602)	Novel Protein sim. GBank gl4082973 gb BAA36204.1 - (AB017138) alpha subunit of malonate decarboxylase [Pseudomonas putida]	UNCLASSIFIED	284906	
1302	95416198 (2603, 2604)			85658542, 285020	
1303	9684121 (2605, 2606)			284908	
1304	78377188 (2607, 2608)		UNCLASSIFIED	284508	
1305	19905889 (2609, 2610)			284566	
1306	13089230 (2611, 2612)	Novel Protein sim. GBank gl3242273 emb CAB07017 - (Z92689) hypothetical protein Rv0236c [Mycobacterium tuberculosis]	UNCLASSIFIED	284636	
1307	82201029 (2613, 2614)	Novel Protein sim. GBank gl1502421 (U59433) - 3-ketoacyl acyl carrier protein reductase [Bacillus subtilis]	UNCLASSIFIED	284907, 284592, 264764	
1308	21428814 (2615, 2616)	Novel Protein sim. GBank gl195819 pir S16298 - ferric enterobactin transport protein fepC - Escherichia coli.	reductase	264555	
1309	79263011 (2617, 2618)	Novel Protein sim. GBank gl195819 pir S16298 - ferric enterobactin transport protein fepC - Escherichia coli.	transport	284906, 18108354	
1310	20466319 (2619, 2620)	Novel Protein sim. GBank gl15459220 emb CAB48893.1 - (AL098837) putative iron-sulfur protein [Streptomyces coelicolor A3(2)]	UNCLASSIFIED	284605	
1311	87813142 (2621, 2622)			35698286, 29331827, 264908, 265008, 284764, 264786, 264886, 21908787, 21908789, 35695917, 284691, 264693	
1312	88061720 (2623, 2624)	Novel Protein sim. GBank gl4455118 gb AAD21084 - (AF125158) zinc finger DNA binding protein 99 [Homo sapiens]	dna_ma_bind	22278995, 22278998, 22278999, 264905, 264908, 265011, 265017, 265019, 284687, 21908788, 265020, 265021, 33657023, 22279002, 264564	
1313	91225458 (2625, 2626)	Novel Protein sim. GBank gl4929733 gb AAD34127.1 AF15189 - (AF15189) CGI-132 protein [Homo sapiens]	ribosomalprot	22278996, 22278999, 264259, 20281099, 29146496, 264508, 264908, 68712502, 60433356, 60433438, 265011, 265017, 264863, 284288, 21908785, 21908787, 29148827, 21908788, 35695917, 265021, 33657023, 33657109, 18108370, 18108377, 35695955, 60432113, 22279000, 264563, 18108390	
1314	56928053 (2627, 2628)			264693	
1315	84357192 (2629, 2630)	Novel Protein sim. GBank gl12589223 (AF026565) - ring finger protein [Mus musculus]	Interleukinrecept	264691	

1316	95361609 (2631, 2632)	Novel Protein sim. GBank glj5689407[dbj BAA82987.1] - (AB028958) KIAA1035 protein [Homo sapiens]		kinase	56182575, 56181696, 20281171, 29331822, 29331824, 60424289, 29331825, 35696052, 52644045, 264591, 60432229, 265018, 265019, 55811150, 56181562, 21906765, 21908787, 21908788, 35695917, 60170815, 33657023, 65274620, 33657109, 35695763, 35695855, 18108387, 87168518, 60432113, 22279002, 264564
1317	88055167 (2633, 2634)	Novel Protein sim. GBank glj4836757[gb AAD30541.1]AF13491 - (AF134918) semaphorin subclass 4 member G [Mus musculus]		UNCLASSIFIED	264093, 264906, 264908, 264369, 264684
1318	95322893 (2635, 2636)	Novel Protein sim. GBank glj4660204[gb AAD27567.1]AF11417 - (AF114171) hypothetical protein [Sorghum bicolor]		UNCLASSIFIED	18108392, 18108348, 265011, 265017, 18108359, 18108362, 56182323, 18108365, 22279000
1319	94238546 (2637, 2638)				264908, 264909, 265008, 265008, 264592, 265019, 264766, 56181562, 18108368, 264628, 264629, 18108377, 264638
1320	86603567 (2639, 2640)	Novel Protein sim. GBank glj4240183[dbj BAA74870.1] - (AB020854) KIAA0847 protein [Homo sapiens]		UNCLASSIFIED	35696286, 55812038, 265016, 21906768, 265020, 263978, 22279002
1321	86676351 (2641, 2642)	Novel Protein sim. GBank glj4886505[emb CAB43377.1] - (AL050276) hypothetical protein [Homo sapiens]	Contains protein domain (PF00651) - BTB/POZ domain	transcript factor	60432048, 29331828, 264907, 264908, 264909, 264910, 55812038, 264601, 264762, 264764, 264766, 264768, 264769, 264628, 18108374, 264634, 264635, 18108385
1322	87755272 (2643, 2644)	Novel Protein sim. GBank glj5282591[emb CAB45738.1] - (AL080143) hypothetical protein [Homo sapiens]	Contains protein domain (PF00096) - dna_ma_bind		29331828, 264908, 265020, 33657023, 264693, 264404
1323	94845931 (2645, 2646)	Novel Protein sim. GBank glj5459516[dbj BAA82407.1] - (AB029821) phosphatidylethanolamine N-methyltransferase [Homo sapiens]	Zinc finger, C2H2 type	synthase	65274572, 56994075, 264259, 29331822, 29331827, 264104, 56182435, 87168474, 18108351, 264288, 21906768, 21906767, 35695917, 265020, 264693, 65274781, 56182323, 18108387
1324	87737614 (2647, 2648)	Novel Protein sim. GBank glj5031717[ref NP_005704.1]pGPBP - goodpasture antigen-binding protein	Contains protein domain (PF01852) - START domain		22278998, 22278998, 29331828, 264905, 264907, 29331830, 264908, 264510, 265008, 264595, 264759, 21906754, 285018, 264288, 264768, 264769, 21906768, 265022, 18108376, 264631, 264632, 264634, 264638, 264638, 264583, 264584, 264585, 264586
1325	94847471 (2649, 2650)	Novel Protein sim. GBank glj3294501 (U64857) - similar to the DPT/Kunitz family of inhibitors; most similar to tissue factor pathway inhibitor precursor [Caenorhabditis elegans]	Contains protein domain (PF00090) - Thrombospondin type 1 domain	protease	35696288, 264905, 264908, 264907, 264908, 264909, 264910, 264593, 33657402, 264758, 85658542, 264760, 264768, 264769, 264681, 35696423
1326	87316289 (2651, 2652)	Novel Protein sim. GBank glj1397275 (U61947) - C06G3.8 gene product [Caenorhabditis elegans]		UNCLASSIFIED	264259, 66712502, 264682, 264683, 264635

1327	95322897 (2653, 2654)	Novel Protein sim. GBank gi 728832 sp P39189 ALU2_HUMAN - IIII ALU SUBFAMILY SB WARNING ENTRY IIII	Contains protein domain (PF00279) - Plant lipid transfer protein family	UNCLASSIFIED	18108398, 22278996, 22278997, 22278999, 284091, 284259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 284105, 264905, 56182435, 264112, 265008, 265009, 21908754, 265010, 265011, 265017, 265019, 264681, 264448, 264784, 264684, 264288, 264685, 264768, 264688, 21908787, 21906769, 28148629, 265020, 284690, 284691, 284692, 284693, 283987, 33657109, 33657182, 27486262, 33657349, 18108370, 18108374, 55810784, 35695855, 264634, 56182323, 83373044, 87188518, 60432113, 22279000, 22279002, 264563, 264567 UNCLASSIFIED
1328	87753493 (2655, 2656)			UNCLASSIFIED	18108351, 264766, 22279002, 284482, 284567
1329	87755276 (2657, 2658)	Novel Protein sim. GBank gi 4678224 gb AAD26969.1 AC00713 - (AC007135) unknown protein [Arabidopsis thaliana]		UNCLASSIFIED	22278996, 29331827, 264684, 264692, 33657109
1330	87727737 (2659, 2660)	Novel Protein sim. GBank gi 437310 (L23504) - nodulin [Medicago truncatula]		UNCLASSIFIED	264259, 29331825, 284512, 265019, 265021, 284555, 264558, 56528486
1331	87376764 (2681, 2682)	Novel Protein sim. GBank gi 4589586 dbj BAA76815.1 - (AB023188) KIAA0971 protein [Homo sapiens]		UNCLASSIFIED	264259, 29331826, 29331827, 35698052, 29331828, 60170831, 264448, 264686, 21908765, 55811957, 265020, 33657023, 33657109, 263973, 55811576, 35698423, 35695855, 56182323
1332	94845937 (2683, 2684)	Novel Protein sim. GBank gi 5459516 dbj BAA82407.1 - (AB029821) phosphatidylethanolamine N-methyltransferase [Homo sapiens]	synthase		65274572, 22278996, 56994075, 22278999, 60432049, 264259, 29331822, 29331828, 60432289, 29331827, 35698052, 52644045, 56182435, 284510, 21908754, 87168559, 265018, 265019, 264448, 264288, 264389, 264688, 21908765, 21908766, 21908787, 21908768, 265020, 265021, 52644150, 33657023, 56182323, 18108387, 60432113, 22279002
1333	88098476 (2685, 2686)	Novel Protein sim. GBank gi 5689527 dbj BAA83047.1 - (AB029018) KIAA1095 protein [Homo sapiens]		UNCLASSIFIED	60432289, 66712502, 264591, 60433358, 60433438, 55812038, 265010, 264639, 56528488
1334	87592388 (2687, 2688)	Novel Protein sim. GBank gi 2662536 (AF036685) - Similar to protein-tyrosine phosphatase [Caenorhabditis elegans]	phosphatase		264905
1335	87644798 (2689, 2690)	Novel Protein sim. GBank gi 4240285 dbj BAA74921.1 - (AB020705) KIAA0888 protein [Homo sapiens]	Contains protein domain (PF00843) - B-box zinc finger.	UNCLASSIFIED	22278998, 22278999, 29331827, 264509, 264511, 265007, 265008, 265009, 60433438, 21908754, 87168559, 265017, 284288, 21908765, 21908767, 21908768, 21908769, 265020, 33657109, 27486264, 18108374, 264556, 264638, 264557, 60170394, 264559, 18108385, 264563

1336	87787890 (2671, 2672)	Novel Protein sim. GBank gi 465445 sp P33485 VNUA_PRVKA - PROBABLE NUCLEAR ANTIGEN			264509, 264905, 264512, 264764, 264683, 264635, 264637
1337	94312042 (2673, 2674)	Novel Protein sim. GBank gi 568947 dbj BAA83019.1 - (AB028990) KIAA1067 protein [Homo sapiens]		UNCLASSIFIED	56182575, 56994075, 22278998, 22278999, 284092, 264259, 60432289, 28331826, 264906, 264908, 264909, 264112, 265008, 265009, 60433356, 55812038, 33657084, 265011, 265017, 265018, 265019, 264682, 264448, 264683, 264369, 264688, 264689, 21908769, 21908769, 265020, 264691, 27488281, 20281089, 18108379, 55811576, 35695855, 56182323, 60432113, 22279002, 264587
1338	80366114 (2675, 2676)			UNCLASSIFIED	28331822, 265010, 264288, 264689, 18108370, 35695855
1339	80249231 (2677, 2678)	Novel Protein sim. GBank gi 1176422 (U43194) - rhophilin [Mus musculus]		UNCLASSIFIED	35696052, 264909, 264688, 264556, 264558
1340	86316311 (2679, 2680)				264905, 264907, 87168559, 264764
1341	86101485 (2681, 2682)				264681, 264685, 264688, 264692
1342	80089017 (2683, 2684)	Novel Protein sim. GBank gi 5019564 emb CAB44507.1 - (AL035942) dJ994E9.5 (hsGM1-17 (novel 7 transmembrane receptor (rhodopsin family) (olfactory receptor like protein)) [Homo sapiens])	Contains protein domain (PF00001) - Im7 7 transmembrane receptor (rhodopsin family)		264629
1343	80082862 (2685, 2686)	Novel Protein sim. GBank gi 4557543 ref NP_001384.1 pECM2 - extracellular matrix protein 2	Contains protein domain (PF00560) - struct Leucine Rich Repeat		264910, 264686, 264534
1344	20562559 (2687, 2688)				263978
1345	91225546 (2689, 2690)	Novel Protein sim. GBank gi 2144101 p p 55210 - Intercarboxylate carrier - rat (fragment)		glycoprotein	264909, 60170394
1346	80255717 (2691, 2692)	Novel Protein sim. GBank gi 3881052 emb CAA19523 - (AL023843) predicted using GeneFinder, similar to serine/threonine kinase; cDNA EST yk246a12.3 comes from this gene; cDNA EST yk396c10.5 comes from this gene; cDNA EST EMBL:M89047 comes from this gene; cDNA EST yk246a12.5 comes...	Contains protein domain (PF00069) - kinase Eukaryotic protein kinase domain		22278998, 264907, 264681, 264685, 264689, 265020, 264693, 22279000, 22279002, 264596
1347	80417393 (2693, 2694)	Novel Protein sim. GBank gi 4504379 ref NP_003658.1 pHG38 - orphan G protein- coupled receptor HG38	Contains protein domain (PF00560) - glycoprotein Leucine Rich Repeat		264906, 264908, 264909, 265006, 264910, 265011, 265017, 264764, 264766, 264787, 264769, 264631, 264634, 264638, 264567, 264486
1348	87352335 (2695, 2696)	Novel Protein sim. GBank gi 3399720 dbj BAA32100 - (AB010999) peptidylarginine deiminase type IV [Rattus norvegicus]		UNCLASSIFIED	264488, 264489, 264508, 264509, 264510, 264511, 264512, 264591, 264592, 264601, 264684, 264685, 264769, 264532, 264534, 264553, 264556, 264557, 264558, 22279002, 264486

1349	91225548 (2697, 2698)	Novel Protein sim. GBank gij2144101[pir]i55210 - lricarboxylate carrier - rat (fragment)		UNCLASSIFIED	52646842, 35698288, 22278998, 22278998, 22278999, 264259, 29331822, 29331824, 35698052, 264509, 264905, 264906, 264907, 264908, 264909, 264511, 265008, 264512, 264910, 60170831, 264591, 60433338, 264757, 21908754, 265017, 265018, 264605, 264760, 264762, 264288, 264786, 264689, 21908785, 21908786, 21908787, 21908788, 55811957, 35695917, 265020, 264534, 264691, 264692, 33657023, 264693, 33657349, 18108374, 18108376, 35696423, 60170394, 22278000, 22278002, 264563, 264584
1350	87093136 (2689, 2700)			UNCLASSIFIED	52646842, 264259, 29331825, 264908, 264511, 264604, 264288, 21908789, 265020, 33657182, 33657349, 18108374, 35695855, 264555, 264556, 18108385, 22278002, 264486
1351	87361327 (2701, 2702)	Novel Protein sim. GBank gij4887239[gb]AAD32246.1] - (AF064584) BAW protein [Fugu rubripes]		UNCLASSIFIED	264906, 264907, 264638
1352	80076386 (2703, 2704)			UNCLASSIFIED	264693, 263981
1353	85345417 (2705, 2706)	Novel Protein sim. GBank gij2144101[pir]i55210 - lricarboxylate carrier - rat (fragment)		UNCLASSIFIED	35696286, 60424269, 264905, 264509, 264906, 264907, 264908, 264909, 264511, 264512, 264910, 264758, 264596, 55811386, 265011, 264605, 55811150, 264762, 264764, 264766, 52644229, 58181562, 35695917, 265022, 33657023, 264693, 35695783, 60431528, 264629, 263978, 35696423, 35695855, 264630, 264634, 264635, 264636, 264637, 264638, 264639, 18108385, 264563, 264564, 264566
1354	85350845 (2707, 2708)	Novel Protein sim. GBank gij4638108[gb]AAD27763.1[AF07703] - (AF077030) hypotheical 43.2 kDa protein [Homo sapiens]		UNCLASSIFIED	22278895, 22278898, 29331826, 264906, 265008, 33657402, 21908754, 265011, 67188559, 264684, 264369, 264769, 264689, 21908785, 21908786, 52644150, 33657023, 264692, 264693, 18108374, 83373044, 87188518, 22278000
1355	88260186 (2709, 2710)	Novel Protein sim. GBank gij1469199[db]BAA09487] - (D50t28) The KIAA0138 gene product is novel. [Homo sapiens]	Contains protein domain (PF00078) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	- dna_ma_bhd	22278898, 22278897, 264259, 60714117, 264511, 21908754, 265010, 264769, 264689, 21908785, 21908786, 21908789, 264532, 27488282, 264629, 264636, 264556, 264638, 264639, 264482, 264484

1356	85313981 (2711, 2712)	Novel Protein sim. GBank gl 113865 (U40342) - ninein [Mus musculus]		struct	18108397, 22278995, 22278996, 22278998, 284094, 28331828, 284905, 285008, 285007, 285008, 285010, 285017, 285018, 285019, 284784, 18108354, 284688, 21908785, 285022, 18108384, 35898423, 83373044, 18108387
1357	88260268 (2713, 2714)	Novel Protein sim. GBank gl 897693[emb]CAA90330] - (Z50028) phosphatidylcholine transfer protein [Bos taurus]		Contains protein domain (PF01852) - START domain	284259, 29331822, 29331825, 284510, 87168559, 285018, 284448, 284288, 21908785, 21908788, 21908788, 285021, 284693, 18108378
1358	38719455 (2715, 2716)	Novel Protein sim. GBank gl 556219 (L36831) - transcription regulator [Mus musculus]			284787
1359	87771843 (2717, 2718)			UNCLASSIFIED	284907, 284809, 284510, 284511, 284512, 18108351, 284784, 284534, 33857023, 18108374, 284634, 284635, 284638, 284639, 18108385, 284488, 284587
1360	87738272 (2719, 2720)	Novel Protein sim. GBank gl 2588282[emb]CAA75612] - (Y15417) acetate-CoA ligase [Coprinus cinereus]		synthase	80432289, 284605
1361	87583527 (2721, 2722)	Novel Protein sim. GBank gl 5689443[db]BAA83005.1] - (AB028976) KIAA1053 protein [Homo sapiens]		Contains protein domain (PF00538) - SAM domain (Sterile alpha motif)	35898288, 22278997, 22278999, 284259, 29331828, 284508, 284509, 284905, 284907, 284908, 285007, 285009, 33109854, 21808754, 87168474, 285011, 284781, 284683, 284288, 284788, 284788, 284689, 21908788, 285020, 285021, 33857023, 55811578, 35898423, 284634, 80432113, 22279002, 284482, 284488
1362	95287861 (2723, 2724)	Novel Protein sim. GBank gl 5689411[db]BAA82989.1] - (AB028980) KIAA1037 protein [Homo sapiens]		Contains protein domain (PF00400) - WD domain, G-beta repeat	56182575, 56181688, 60432049, 284259, 29331822, 56182181, 29331827, 35898052, 29331828, 284905, 284908, 284908, 284595, 55812038, 85858542, 55811150, 284681, 284288, 284389, 56181582, 60431528, 55810784, 35898423, 60431850, 284559
1363	85758476 (2725, 2726)	Novel Protein sim. GBank gl 130494 (U35776) - ADP-ribosylation factor 1-directed GTPase activating protein [Rattus norvegicus]		Contains protein domain (PF01412) - Putative GTP-ase activating protein for Arf	284488, 29331826, 284607, 284687, 284689, 284693
1364	88178488 (2727, 2728)			UNCLASSIFIED	
1365	83003108 (2729, 2730)	Novel Protein sim. GBank gl 4589562[db]BAA76803.1] - (AB023176) KIAA0959 protein [Homo sapiens]		oncogene	60432289, 60433358, 60433438, 87168559, 284603, 18108351, 21808788, 35898423, 80432113
1366	87003262 (2731, 2732)	Novel Protein sim. GBank gl 1084944[db]J54495 - hypothetical protein YPR021c - yeast (Saccharomyces cerevisiae)		transport	284766

1367	87721210 (2733, 2734)	Novel Protein sim. GBank gij4894088[emb]CAB43240.11 - (AL050019) hypothetical protein [Homo sapiens]	Contains protein domain (PF01342) - SAND domain	UNCLASSIFIED	284488, 52846842, 52846365, 22278995, 59994075, 35898288, 22278996, 22278998, 22278999, 284259, 29331824, 66714117, 29331825, 60432289, 35896052, 264905, 284907, 284908, 284909, 284510, 285008, 285007, 284512, 284910, 285009, 60170831, 33657402, 55812038, 21908754, 285011, 87168559, 285017, 285019, 18108351, 284448, 284682, 284883, 284288, 284389, 284888, 284787, 284689, 21908785, 21908768, 21908768, 21908769, 55811957, 265020, 285021, 285022, 284534, 60170815, 284690, 284691, 18108382, 33657023, 33857109, 33857349, 284628, 18108370, 18108374, 18108378, 55811576, 35898423, 35898555, 284635, 284555, 284637, 284558, 52844332, 60170394, 284558, 18108381, 18108385, 58528486, 22279000, 284583, 284587
1368	94320078 (2735, 2736)	Novel Protein sim. GBank gij484561[sp]P35289[R815_RAT RAS-RELATED PROTEIN RAB-15	Contains protein domain (PF00071) - Ras family	oncogene	284259, 29331822, 29331826, 60432289, 29331827, 35896052, 284508, 284905, 284908, 284908, 284909, 284510, 285007, 284910, 60433438, 284758, 85658542, 87168559, 284600, 284601, 284780, 284784, 264765, 284768, 52844229, 284689, 35895917, 285020, 285021, 284631, 284632, 284634, 284637, 52844332, 284558, 284839, 83373044, 284583, 284588, 284488, 284587, 265008, 60432228, 60433358, 33657084, 21908784, 21908789, 284555, 284638, 284559, 284587
1369	86634033 (2737, 2738)	Novel Protein sim. GBank gij2062702 (U90550) - butyrophilin [Homo sapiens]		UNCLASSIFIED	
1370	95336910 (2739, 2740)	Novel Protein sim. GBank gij5031823[ref]NP_005823.1pKCNM - potassium large conductance calcium-activated channel, subfamily M, beta member 2		potassium_channel	22278998, 284259, 29331822, 29331824, 29331826, 29331827, 35896052, 284508, 284509, 284905, 284908, 284907, 284908, 284909, 284510, 284511, 284512, 284758, 285011, 285019, 284764, 284768, 284768, 21908767, 35895917, 18108362, 35898423, 284832, 284835, 284638, 284555, 284638, 284558, 284639, 18108385, 65274727, 284404, 284583, 284588, 284488
1371	95336512 (2741, 2742)	Novel Protein sim. GBank gij5032203[ref]NP_005714.1pTSPA - tetraspan 5	Contains protein domain (PF00033) - 4 transmembrane segments integral membrane proteins	glycoprotein	22278998, 284259, 29331822, 29331824, 29331825, 29331826, 29331827, 284905, 284509, 29331830, 66712502, 285008, 285009, 284758, 33657084, 85658542, 285010, 285018, 285019, 284762, 284448, 35895917, 33657109, 33657182, 33657349, 35895855, 284558, 22279002, 284583

1372	80248517 (2743, 2744)	Novel Protein sim. GBank gll840708[pibjBAA0334] - (D50885) trans-sialidase [Trypanosoma cruzi]		collagen	263978	264789, 21908765, 21908767, 22278999, 264891, 264910, 55812038, 265010, 264681, 264684
1373	80499421 (2745, 2746)			UNCLASSIFIED		264907, 264510, 285007, 264512, 265008, 60432229, 264689, 65274781, 264555, 264556, 264557, 83373044, 60432113
1374	95087038 (2747, 2748)	Novel Protein sim. GBank gll111876[pir]JC1241 - beta-interferon-induced protein - rat		interferon		52844507, 52845156, 52846842, 52846365, 56162375, 56161688, 22278998, 56994075, 35898286, 22278997, 22278998, 22278999, 264259, 29331822, 52845080, 29331824, 29331825, 60432289, 29331826, 29331827, 29331828, 35896052, 33856970, 264805, 264509, 264908, 264807, 264908, 29331830, 264809, 265008, 264511, 265007, 265008, 285009, 33857402, 60433356, 52646317, 33109954, 33657084, 52844298, 87168474, 87168559, 264600, 265017, 265018, 285019, 55811150, 18108351, 264448, 264784, 264288, 264369, 264766, 52644229, 21908765, 21908768, 21908767, 21908768, 21908769, 55811957, 35895917, 265020, 285021, 52644150, 33657023, 264893, 65274620, 52845128, 33857109, 27488261, 33657349, 27488265, 35895763, 18108376, 55810784, 35898423, 35895855, 264630, 264631, 264634, 264636, 264555, 264638, 18108385, 87168516, 60432113, 22278000, 22279002, 264563, 264584, 264566, 264567
1375	94236942 (2749, 2750)	Novel Protein sim. GBank gll5648176[pib]AAD03500.2] - (AF051155) G beta-like protein GBL [Rattus norvegicus]	Contains protein domain (PF00400) - WD domain, G-beta repeat	kinase		264788, 264768, 35895917, 22278997, 264691, 264259, 29331822, 264693, 35898052, 264508, 264509, 264905, 264906, 264628, 264908, 264828, 18108372, 264809, 264510, 264511, 264512, 285008, 264630, 264631, 264910, 264832, 264634, 264635, 264636, 264591, 264592, 264637, 264638, 264558, 264639, 33857402, 264595, 18108385, 56526488, 265010, 265011, 264800, 264563, 264762, 264564, 264565, 264764, 264488, 264768
1376	87389050 (2751, 2752)	Novel Protein sim. GBank gll136350[sp]P28988[VGLX_HSVB - GLYCOPROTEIN X PRECURSOR		UNCLASSIFIED		264788, 264768, 35895917, 22278997, 264691, 264259, 29331822, 264693, 35898052, 264508, 264509, 264905, 264906, 264628, 264908, 264828, 18108372, 264809, 264510, 264511, 264512, 285008, 264630, 264631, 264910, 264832, 264634, 264635, 264636, 264591, 264592, 264637, 264638, 264558, 264639, 33857402, 264595, 18108385, 56526488, 265010, 265011, 264800, 264563, 264762, 264564, 264565, 264764, 264488, 264768
1377	86964242 (2753, 2754)	Novel Protein sim. GBank gll1663848 (U75321) - chromaffin granule ATPase II homolog [Mus musculus]	Contains protein domain (PF00122) - E1-E2 ATPase	ATPase-associated		264788, 264768, 35895917, 22278997, 264691, 264259, 29331822, 264693, 35898052, 264508, 264509, 264905, 264906, 264628, 264908, 264828, 18108372, 264809, 264510, 264511, 264512, 285008, 264630, 264631, 264910, 264832, 264634, 264635, 264636, 264591, 264592, 264637, 264638, 264558, 264639, 33857402, 264595, 18108385, 56526488, 265010, 265011, 264800, 264563, 264762, 264564, 264565, 264764, 264488, 264768

1378	87595071 (2755, 2756)	Novel Protein sim. GBank gll4107015[db][BAA36293] - (AB001772) PEM-5 [Ciona savignyi]			22278994, 22278998, 264093, 264094, 264259, 29331824, 29331827, 265009, 265018, 265019, 18108351, 264369, 264288, 29148627, 55811957, 264691, 18108368, 33657109, 18108368, 264635, 263981, 18108385
1379	85879344 (2757, 2758)	Novel Protein sim. GBank gll3252872 (AF035620) - BRCA1-associated protein 2 [Homo sapiens]			265020
1380	87627862 (2759, 2760)	Novel Protein sim. GBank gll4637737[gb][AAD30662.1] - (AF098834) germ cell specific Y-box binding protein [Homo sapiens]		nuc1_rept	264510, 264512, 265009, 264288, 264564
1381	88178656 (2761, 2762)	Novel Protein sim. GBank gll4731580[gb][AAD28508.1] (AF12538) L82A [Drosophila melanogaster]		UNCLASSIFIED	87168559, 265017, 264628, 22279002
1382	94847576 (2763, 2764)		Contains protein domain (PF00583) - Acetyltransferase (GNAT) family	UNCLASSIFIED	22278994, 22278997, 22278999, 29331822, 29331824, 29331826, 265007, 60432229, 60433356, 85858542, 265017, 265018, 264685, 264768, 21908768, 35695917, 33657023, 27486281, 27486282, 35895763, 35695855, 87168518, 22279002
1383	87860598 (2765, 2766)			UNCLASSIFIED	18108386, 264692
1384	86915895 (2767, 2768)			UNCLASSIFIED	264488, 264508, 264509, 264905, 264908, 264908, 264909, 264511, 264512, 264910, 264760, 18108351, 264768, 264769, 35695855, 264630, 264638, 264555, 264638, 264483, 264584, 264486
1385	86378788 (2769, 2770)	Novel Protein sim. GBank gll2384732 (AF015911) - NAC-1 protein [Rattus norvegicus]		UNCLASSIFIED	35696052, 55811386, 264688, 21906765, 265020, 33657023, 18108385
1386	91013049 (2771, 2772)	Novel Protein sim. GBank gll2384910 (AF022982) - contains similarity to the A-type potassium current class of channel proteins [Caenorhabditis elegans]		Inf	60432289, 29331828, 264908, 264907, 56182435, 285011, 264681, 60170815, 33657023, 83373044, 264566
1387	87797958 (2773, 2774)	Novel Protein sim. GBank gll4160304[emb][CAA10600] - (AJ132192) HSI binding protein 3 [Mus musculus]		UNCLASSIFIED	264591
1388	95101652 (2775, 2776)	Novel Protein sim. GBank gll4695164[gb][AAD32753.1] (AC007231) putative disease resistance protein [Arabidopsis thaliana]		glycoprotein	65274572, 22278999, 264259, 29331826, 29331827, 35696052, 264509, 264907, 264908, 264909, 265006, 265008, 60170831, 33657402, 60433438, 264598, 21906754, 87168559, 264600, 265017, 264683, 18108354, 52844229, 21906765, 21906768, 21906767, 21906768, 21906769, 265021, 264692, 33657023, 33657109, 35895855, 264558, 60170394, 83373044, 22279000

1389	91256016 (2777, 2778)	Novel Protein sim. GBank gij5689387[dbj BAA82977.1] - (AB028948) KIAA1025 protein [Homo sapiens]	Contains protein domain (PF00641) - Zn-finger in Ran binding protein and others.	UNCLASSIFIED	65274572, 22278999, 264259, 29331822, 29331824, 29331825, 60432289, 29331828, 29146489, 264908, 68712502, 55812038, 265017, 265018, 265019, 18108351, 284389, 21908765, 21908768, 21908767, 21908768, 265020, 265021, 284692, 33657023, 33657349, 18108370, 18108374, 55811576, 284555, 264558, 264557, 60170394, 83373044, 22278000, 264563, 264564, 52645156, 52646365, 264259, 52645080, 29331825, 29331826, 264908, 52644045, 265009, 33657084, 52644296, 87168474, 87168559, 265017, 265018, 264760, 264682, 264288, 264688, 264687, 56181562, 52644229, 21908765, 21908769, 35695917, 265020, 265021, 60170615, 52644150, 33657023, 27486281, 27486284, 65274791, 264631, 264555, 52644332, 87168518, 22279000, 264567
1390	94111916 (2779, 2780)	Novel Protein sim. GBank gij3702295 (AC005783) - R33083_1 [Homo sapiens]	peptidase		29331826, 29331828, 29331830, 284448, 284288, 33657023, 18108365, 264555, 284558, 83373044
1391	91227345 (2781, 2782)	Novel Protein sim. GBank gij1346910[sp P28650 PUA1_MOUSE - ADENYLOSUCCINATE SYNTHETASE, MUSCLE ISOZYME (IMP--ASPARTATE LIGASE)	Contains protein domain (PF00709) - Adenylosuccinate synthetase		52646842, 65274572, 22278994, 22278995, 35696286, 56994075, 22278997, 22278998, 22278999, 60432049, 264259, 52645080, 29331822, 29331824, 60432289, 29331826, 29331827, 35696052, 29331828, 33656970, 284907, 52644045, 265006, 265007, 285008, 60431735, 60433356, 52646317, 55811388, 52644286, 265010, 87168559, 265017, 284604, 265018, 265019, 264448, 264288, 264369, 264768, 21906764, 21906767, 35695917, 265020, 265021, 33657109, 52645129, 27486261, 27488262, 27488265, 33657349, 35695763, 18108370, 18108374, 18108376, 55811576, 35696423, 35696855, 264636, 52644332, 18108382, 18108385, 87168518, 60432113, 22278000, 264484, 284566, 18108391, 264763, 264631
1392	94311097 (2783, 2784)	Novel Protein sim. GBank gij728286 (U22394) - mSin3A [Mus musculus]			
1393	80409472 (2785, 2786)		Contains protein domain (PF00560) - Leucine Rich Repeat	UNCLASSIFIED	
1394	15028819 (2787, 2788)			UNCLASSIFIED	284629
1395	95361471 (2789, 2790)	Novel Protein sim. GBank gij2274845[dbj BAA21534] - (D88481) N-WASP [Rattus rattus]		UNCLASSIFIED	265009, 18108381

1396	95363233 (2791, 2792)	Novel Protein sim. GBank gi 2135904 pr j 54810 - pHL EIF1 - human			22278997, 22278999, 264259, 29331825, 60432289, 29331828, 29148498, 29148499, 264907, 264908, 29331830, 264909, 265006, 265007, 265008, 265009, 60433356, 265010, 264802, 265017, 265018, 265019, 18108354, 52844229, 18108358, 21908787, 29148827, 21908788, 21908789, 29148629, 29148784, 265021, 265022, 18108368, 18108374, 56182323, 18108385, 264563, 264567, 35696286, 264907, 66712502, 264510, 35695917, 264692, 264693, 35696423
1397	87631317 (2793, 2794)			UNCLASSIFIED	
1398	91233667 (2795, 2799)	Novel Protein sim. GBank gi 5420389 emb CAB46680.1 - (AJ243460) proteophosphoglycan [Leishmania major]			264259, 29331822, 29331824, 29331825, 29331827, 35696052, 33656970, 87168474, 265018, 265019, 264682, 264768, 21908787, 265020, 33657023, 27486261, 55811576, 264632, 264639, 83373044, 87168518, 22278002
1399	87631076 (2797, 2798)	Novel Protein sim. GBank gi 2496887 sp Q09233 YQ22_CAEEL - HYPOTHETICAL 32.0 KD PROTEIN C09F5.2 IN CHROMOSOME III		UNCLASSIFIED	264768, 18108370, 264555, 264557
1400	95419084 (2799, 2800)	Novel Protein sim. GBank gi 283920 pr j 527839 - tensin - chicken	Contains protein domain (PF00017) - Src homology domain 2	UNCLASSIFIED	56182575, 22278994, 22278997, 264259, 29331822, 29331825, 29331828, 29331827, 29331828, 264908, 56182435, 264112, 265009, 265011, 265017, 265018, 265019, 264760, 264762, 264765, 264288, 264885, 264687, 56181562, 264769, 21908788, 21908787, 55811957, 264691, 264692, 264628, 264629, 55811576, 264634, 264555, 264637, 264557, 264638, 18108381, 264558, 18108384, 60432113, 22278000
1401	91226379 (2801, 2802)	Novel Protein sim. GBank gi 3256185 emb CAA15485 - (AL008635) dJ510H16.1 [Homo sapiens]	Contains protein domain (PF00790) - VHS domain	UNCLASSIFIED	65274572, 60432289, 264909, 264758, 264768, 21908769, 22278002
1402	95361475 (2803, 2804)	Novel Protein sim. GBank gi 1515427 (U57523) - nel homolog [Homo sapiens]	Contains protein domain (PF00008) - EGF-like domain	IgI	264905, 264907, 264908, 264909, 264112, 264693, 33657109, 264634
1403	94147933 (2805, 2806)	Novel Protein sim. GBank gi 5262815 emb CAB45747.1 - (AL080156) hypothetical protein [Homo sapiens]			65274572, 66712502, 265017, 264448, 264288, 21908765, 21908769, 264693, 55811576, 65274791, 60432113
1404	90935393 (2807, 2808)			UNCLASSIFIED	65274572, 22278998, 29331822, 29331828, 66712502, 265008, 60433438, 265017, 264693, 18108385

1405	95095068 (2809, 2810)	Novel Protein sim. GBank gij854065[emb]CAA58337] - (X83413) U88 [Human herpesvirus 6]			264488, 56994075, 35698286, 29331822, 29331824, 29331828, 29331828, 35698052, 264508, 264908, 264907, 264908, 284510, 264511, 284910, 33857402, 284594, 284758, 264800, 264604, 264762, 18108351, 264764, 33857023, 33857109, 264628, 264634, 83373044, 22278002, 264563, 264482, 264488, 264587
1406	87612368 (2811, 2812)	Novel Protein sim. GBank gij624076[gb]AAC98425.1] - (U42580) contains Pro-rich Px motifs: SPKPP (20X), PEPPA (9X); similar to soybean pro-rich cell wall protein, corresponds to Swiss-Prol Accession Number P13993 [Paramedum bursaria Chlorella virus 1]	collagen		264907, 264605
1407	94129872 (2813, 2814)	Novel Protein sim. GBank gij2827886 (AF015037) - endooligopeptidase A related protein; EOPA related protein [Oryctolagus cuniculus]	UNCLASSIFIED		35698286, 22278999, 264084, 264259, 66714117, 29331828, 29331827, 29331828, 29146498, 284107, 264908, 285008, 285008, 284910, 60433438, 285011, 285017, 18108351, 284448, 264288, 264686, 21908765, 21908769, 264692, 33857109, 18108370, 264628, 263972, 18108374, 35698423, 55811576, 264631, 284557, 284558, 83373044, 18108385, 87168518, 80432113, 22279002
1408	95361477 (2815, 2816)	Novel Protein sim. GBank gij2664953 (AF030001) - unknown [Mus musculus]	Contains protein domain (PF00008) - oncogene EGF-like domain		264488, 264489, 35698286, 264109, 264508, 264905, 264509, 264908, 264907, 264908, 264909, 265008, 265009, 264910, 33657402, 284757, 264758, 285011, 284601, 285017, 284760, 264762, 264683, 284685, 284766, 284687, 264689, 21908767, 265021, 284690, 264691, 33657023, 264692, 264693, 33857109, 264628, 264629, 35698423, 35695855, 264631, 264632, 264634, 264635, 284555, 284638, 264637, 264638, 56182323, 284639, 264563, 264564, 264565, 264566, 284587
1409	86644385 (2817, 2818)	Novel Protein sim. GBank gij2682185[db]BAA23714] - (AB007902) HH0712 cDNA clone for KIAA0442 has a 574-bp insertion at position 1474 of the sequence of KIAA0442. [Homo sapiens]			284693
1410	86612587 (2819, 2820)	Novel Protein sim. GBank gij2493780[sp]Q80994]ACR3_MOUSE - 30 KD ADIPOCYTE COMPLEMENT-RELATED PROTEIN PRECURSOR (ACRP30) (ADIPOCYTE SPECIFIC PROTEIN ADIPOQ)	Contains protein domain (PF00386) - complement C1q domain		29331826, 264112, 284512, 265009, 265010, 264601, 264686, 284768, 21908767, 263974, 284631, 284568

1411	87818641 (2821, 2822)	Novel Protein sim. GBank glij3123155jip91343YM3M_CAEEL - HYPOTHETICAL 49.0 KD TRP-ASP REPEATS CONTAINING PROTEIN F55F8.5 IN CHROMOSOME I	Contains protein domain (PF00400) - kinase receptor WD domain, G-beta repeat	22278995, 22278997, 22278999, 29331822, 60432289, 29331828, 284907, 265017, 265018, 284682, 21906787, 21906788, 21906789, 265020, 284690, 284691, 33657023, 33657109, 27486284, 284628, 283972, 284634, 284558, 18108385 264757
1412	84390919 (2823, 2824)		UNCLASSIFIED	
1413	95416559 (2825, 2826)	Novel Protein sim. GBank glij3879121jemb CAA94370 - (270310) predicted using Genefinder; Similarity to Mouse ankyrin (PIR Acc. No. S37771); cDNA EST EMBL: T01923 comes from this gene; cDNA EST EMBL: D32335 comes from this gene; cDNA EST EMBL: D32723 comes from this gene; cDNA EST E...	Contains protein domain (PF00023) - Ank repeat	56994075, 29331822, 35695052, 29331828, 29331830, 284809, 52644045, 284510, 52644296, 85658542, 87168474, 265017, 265018, 284681, 284687, 21906788, 35695917, 265020, 52644150, 284692, 263967, 27488264, 35695763, 284639, 18108387, 284568 264682, 284683, 265022, 264636
1414	94675660 (2827, 2828)	Novel Protein sim. GBank glij3252981 (AF0888921) - Ras- binding protein SUR-8 [Mus musculus]	Contains protein domain (PF00560) - Leucine Rich Repeat	
1415	94326948 (2829, 2830)	Novel Protein sim. GBank glij1871187 (U90439) - unknown protein [Arabidopsis thaliana]		52646365, 58182575, 22278994, 22278995, 56994075, 22278998, 22278997, 22278998, 22278999, 284259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 29146498, 66712502, 29331830, 52644045, 264113, 264511, 33657402, 264757, 21906754, 55811386, 265017, 285018, 265019, 264761, 284683, 284369, 264288, 264686, 284689, 21906766, 21906787, 29148827, 21906769, 55811957, 265020, 285021, 284690, 33657023, 65274620, 52645129, 27488282, 27486284, 60431528, 284629, 35695855, 56182323, 264559, 60432113, 284404, 22279002, 284482

1418	84325977 (2831, 2832)	Novel Protein sim. GBank gi15106557 gb AAD39749.1 AF12305 - (AF123052) MLL sepin-like fusion protein [Homo sapiens]	Contains protein domain (PF00735) - struct Cell division protein	18108392, 18108394, 18108397, 18108398, 22278995, 22278998, 35886286, 22278997, 22278998, 22278999, 29331822, 35698052, 29331828, 29146498, 264805, 264907, 264908, 264828, 264909, 264113, 265008, 285007, 265008, 265009, 60170831, 264595, 18108348, 21906754, 87168474, 265010, 265011, 87168559, 265017, 265018, 265019, 264762, 18108351, 264681, 264763, 264682, 264683, 264766, 52644229, 264686, 264689, 21906785, 21906786, 21906787, 29146627, 21906788, 55811957, 29146629, 265020, 52644150, 18108361, 33657023, 18108362, 18108368, 264628, 18108370, 264629, 18108374, 18108379, 55811576, 65274791, 264634, 264636, 56182323, 18108381, 60170394, 18108385, 56526486, 87168518, 22279000
1417	87826663 (2833, 2834)	Novel Protein sim. GBank gi14958835 gb BAA78095.1 - (AB027570) suppressor of potassium transport defect 3 [Rattus norvegicus]	ATPase-associated	264107, 264448
1418	87594276 (2835, 2836)		UNCLASSIFIED	264259, 264908, 265010, 52644228, 21906764, 21906768, 264690, 264639, 18108388
1419	87757168 (2837, 2838)	Novel Protein sim. GBank gi12072294 (U95097) - mitotic phosphoprotein 43 [Xenopus laevis]	struct	264259, 60432289, 265006, 87168474, 264288
1420	87298628 (2839, 2840)	Novel Protein sim. GBank gi15174421 ref NP_086023.1 pCPNE - copine VI (neuronal) (Z88635) Similarity to Yeast uridine kinase (SW:URK1_YEAST); cDNA EST EMBL:Z14695 comes from this gene; cDNA EST CEMSE17F comes from this gene; cDNA EST EMBL:D67355 comes from this gene; cDNA EST YK209h1.5 comes from this gene...	ATPase-associated	29331824, 265007, 264563
1421	94746986 (2841, 2842)	Novel Protein sim. GBank gi13876090 emb CAA93459.1 - (Z88635) Similarity to Yeast uridine kinase (SW:URK1_YEAST); cDNA EST EMBL:Z14695 comes from this gene; cDNA EST CEMSE17F comes from this gene; cDNA EST EMBL:D67355 comes from this gene; cDNA EST YK209h1.5 comes from this gene...	Kinase	18108358, 18108396, 18108397, 21906768, 18108398, 21906767, 56182575, 21906768, 21906769, 56181686, 55811957, 35695917, 35886286, 22278996, 22278997, 22278998, 22278999, 265021, 265022, 60170615, 264259, 33657023, 29331822, 56182181, 29331824, 66714117, 29331825, 33657109, 29331826, 27486261, 29331828, 35886052, 33657349, 264905, 264509, 20281149, 18108370, 264907, 60431528, 66712502, 263972, 55811578, 35886423, 35695855, 264512, 265007, 60431850, 60432228, 60431735, 56182323, 264558, 60170394, 83373044, 55812038, 264758, 18108385, 21906754, 55811386, 87168518, 87168559, 60432113, 265017, 265018, 265019, 22279002, 55811150, 264563, 264682, 264763, 264448, 264586, 264486, 18108391

1422	88178777 (2843, 2844)	Novel Protein sim. GBank gi14503939[refNP_000928.1]pPOLR - polymerase (RNA) II (DNA directed) polypeptide A (220kD)			mapolymerase	56994075, 35698288, 87168559, 55811957, 55811578, 284555, 284557, 87168518
1423	86997762 (2845, 2846)				UNCLASSIFIED	284686, 284489, 284692, 284594, 284603, 285018, 284908
1424	95201610 (2847, 2848)	Novel Protein sim. GBank gi1437181 (U02289) - GTPase- activating protein [Caenorhabditis elegans]	Contains protein domain (PF00620) - struct RhoGAP domain			29331822, 29331825, 29331827, 29148498, 284905, 284908, 284908, 284909, 285007, 284910, 285009, 33109954, 285010, 87168559, 285019, 284786, 284887, 21906765, 21906766, 21906767, 21906768, 29148627, 55811957, 29148629, 285021, 284691, 284692, 56528486, 22278002, 284583
1425	21662314 (2849, 2850)	Novel Protein sim. GBank gi100798[prj1S14959 - proline- rich protein - wheat]			UNCLASSIFIED	285007, 284558
1426	94322115 (2851, 2852)	Novel Protein sim. GBank gi12078441 (U56964) - weak similarity to S. cerevisiae intracellular protein transport protein US1 (SP:P25386) [Caenorhabditis elegans]			UNCLASSIFIED	284488, 80424179, 35698288, 22278997, 22278998, 22278999, 284259, 60432049, 29331822, 29331825, 29331827, 56182435, 284910, 60433356, 60433438, 21906754, 285018, 284288, 21906765, 21906766, 21906767, 21906768, 285020, 285022, 33657109, 18108370, 18108376, 284558, 83373044, 18108385, 56528486, 22278002, 264482
1427	91227510 (2853, 2854)	Novel Protein sim. GBank gi15816074[gb AAD45816.1]AF08194 - (AF081943) protate- derived STE20-like kinase PSK [Homo sapiens]	Contains protein domain (PF00089) - kinase Eukaryotic protein kinase domain			22278994, 56994075, 22278997, 29331828, 29331830, 284828, 285006, 285007, 285008, 285009, 284883, 284288, 18108354, 21906765, 21906768, 29148629, 33657023, 18108374, 35698555, 83373044, 22278002, 284584
1428	94323008 (2855, 2856)	Novel Protein sim. GBank gi138350[sp P28986]VGLX_HSVB - GLYCOPROTEIN X PRECURSOR			glycoprotein	56181686, 284259, 284907, 285007, 285009, 284595, 285010, 284888, 65274620, 284829, 65274791, 22278002, 284566
1429	87886889 (2857, 2858)				UNCLASSIFIED	284112, 284595, 285017, 285019, 21906785, 283977, 284555
1430	94735021 (2859, 2860)	Novel Protein sim. GBank gi1181619[dbj BAA11565] - (DB2384) a variant of TSC-22 [Gallus gallus]				284094, 29331824, 284591, 284593, 285018, 284681, 21906765, 21906767, 65274620, 55811576, 284639, 97168518, 22278002
1431	80429081 (2861, 2862)	Novel Protein sim. GBank gi15420389[embj CAB46880.1] - (AJ243460) proteophosphoglycan [Leishmania major]			UNCLASSIFIED	284768, 285020, 284691, 284693, 284638
1432	87463004 (2863, 2864)	Novel Protein sim. GBank gi1414797 (L18986) - Pyruvate dehydrogenase phosphatase [Bos taurus]			phosphatase	18108394, 29148498, 285007, 60433438, 284763, 29148629, 283969
1433	87605403 (2865, 2866)	Novel Protein sim. GBank gi12460316 (AF022147) - uterus- ovary specific putative transmembrane protein [Rattus norvegicus]	Contains protein domain (PF00100) - Zona pellucida-like domain		UNCLASSIFIED	284259, 284510, 284591, 284603, 284565
1434	85113730 (2867, 2868)				UNCLASSIFIED	284882, 284691

1435	84708213 (2869, 2870)	Novel Protein sim. GBank gi 3970850 dbj BAA34789.1 - (AB015330) HRIHFB2007 [Homo sapiens]			transcript factor	22278997, 22278998, 22278999, 60432049, 284259, 29331822, 29331824, 68714117, 29331826, 33656970, 264508, 264905, 68712502, 29331830, 264909, 285007, 265008, 284910, 285009, 60433358, 60433438, 264586, 21906754, 265010, 265017, 265018, 265019, 18108351, 264762, 264448, 284288, 264769, 21908787, 21908768, 21906769, 265020, 265021, 285022, 284680, 264691, 33657109, 264628, 18108374, 18108376, 55811576, 264636, 60170394, 58182323, 264559, 83373044, 87168518, 60432113, 22279000, 22279002, 264563, 264482, 264585, 263978, 284557, 264559
1438	86635024 (2871, 2872)	Novel Protein sim. GBank gi 3183977 emb CAA39515 - (X58044) protein HIFC [Mus musculus]			UNCLASSIFIED	
1437	87631082 (2873, 2874)	Novel Protein sim. GBank gi 2486887 sp Q09232 YQ22, CAEL - HYPOTHETICAL 32.0 KD PROTEIN C09F5.2 IN CHROMOSOME III			UNCLASSIFIED	22278997, 66714117, 29331826, 264907, 56182435, 285009, 18108351, 284692, 284693
1438	85544280 (2875, 2876)	Novel Protein sim. GBank gi 1905908 (AD000092) - hypothetical human serine-threonine protein kinase R31240.1 [Homo sapiens]	Contains protein domain (PF00595) - PDZ domain (Also known as DHR or GLGF).		kinase	284488, 264508, 264908, 264909, 264757, 284600, 264601, 264605, 284788, 264789, 284690, 35898423, 264558, 284583, 264566, 264489, 18108394, 65274572, 56182375, 22278994, 22278998, 35696286, 22278997, 22278998, 22278999, 264259, 60432049, 29331822, 29331824, 60432289, 29331826, 29331827, 29331828, 35896052, 33856970, 264107, 264508, 264509, 264907, 66712502, 29331830, 58182435, 264511, 265008, 285007, 265009, 60432229, 60433438, 264585, 55812038, 55811388, 285011, 265017, 265018, 265019, 18108351, 264448, 18108354, 264288, 18108355, 284767, 21908769, 21908766, 21906767, 21906768, 21908769, 55811957, 35895917, 285020, 265021, 265022, 33657109, 18108370, 18108374, 55810764, 35695855, 284634, 284636, 56182323, 83373044, 18108387, 87168518, 60432113, 22279000, 284488
1439	91231894 (2877, 2878)	Novel Protein sim. GBank gi 3876299 emb CAA94892 - (Z71180) similar to BPTIKUNITZ inhibitor domain; cDNA EST EMBL:D88293 comes from this gene; cDNA EST yk448h4.5 comes from this gene; cDNA EST yk249e6.5 comes from this gene; cDNA EST yk448h4.3 comes from this gene [Caenorhabditis...]	Contains protein domain (PF00450) - Serine carboxypeptidase		calthepsin	22278994, 22278998, 35696286, 22278997, 22278998, 22278999, 264259, 60432049, 29331822, 29331824, 60432289, 29331826, 29331827, 29331828, 35896052, 33856970, 264107, 264508, 264509, 264907, 66712502, 29331830, 58182435, 264511, 265008, 285007, 265009, 60432229, 60433438, 264585, 55812038, 55811388, 285011, 265017, 265018, 265019, 18108351, 264448, 18108354, 264288, 18108355, 284767, 21908769, 21908766, 21906767, 21906768, 21908769, 55811957, 35895917, 285020, 265021, 265022, 33657109, 18108370, 18108374, 55810764, 35695855, 284634, 284636, 56182323, 83373044, 18108387, 87168518, 60432113, 22279000, 284488
1440	87423643 (2879, 2880)	Novel Protein sim. GBank gi 2682165 dbj BAA23714 - (AB007802) HH0712 cDNA clone for KIAA0442 has a 574-bp insertion at position 1474 of the sequence of KIAA0442. [Homo sapiens]			UNCLASSIFIED	284887, 264259, 284906, 264907, 264908, 284909, 285008, 265010, 265017, 265018, 285019, 18108351, 264369, 265020, 33657023, 33657109, 60431528, 55811576, 284635

1441	95317662 (2881, 2882)	Novel Protein sim. GBank gl 4483956 emb CAB11123.2 - (Z88551) predicted using hexExon; MAL3P6.28 (PFC0845c). Hypothetical protein, len: 167 aa; Similarity to model organism hypothetical proteins (C.elegans, D.melanogaster, S.cerevisiae & S.pombe). C.elegans protein ZK287.5 (TR....	Contains protein domain (PF00846) - F-box domain.	18108392, 264488, 263994, 264489, 56182575, 22278994, 22278995, 58994075, 35698286, 22278997, 22278999, 60432049, 264259, 29331822, 29331824, 56182181, 66714117, 29331826, 29331827, 29331828, 35698052, 29146498, 264508, 264905, 264509, 264908, 264907, 264908, 264909, 58182435, 264510, 264511, 264512, 264910, 264592, 264593, 33657402, 60433438, 264595, 264758, 21906754, 85658542, 87168474, 265010, 87168559, 264600, 264602, 265017, 264604, 265018, 264605, 265019, 264760, 264761, 264762, 264681, 264448, 264764, 264683, 264288, 264768, 264768, 264769, 52644229, 264689, 21908765, 21908766, 21908767, 21908768, 21908769, 55811957, 35695917, 265020, 265021, 60170815, 52644150, 264691, 264692, 33657023, 65274620, 33657109, 35695763, 264628, 18108370, 264629, 18108374, 55811576, 35698423, 65274781, 35695855, 264631, 264634, 264635, 264636, 264637, 264638, 56182323, 264558, 60170394, 264639, 264559, 83373044, 18108385, 18108388, 56526486, 22278000, 22278002, 264583, 264483, 264564, 264568, 264567, 264486
1442	83367491 (2883, 2884)	Novel Protein sim. GBank gl 5103027 dbj BAA78765.1 - (AB023419) mSox7 [Mus musculus]	transcription factor	264906, 265007, 264693, 264558
1443	87109935 (2885, 2886)	Novel Protein sim. GBank gl 4887228 gb AAD32244.1 AF15075 - (AF150755) microtubule-actin crosslinking factor [Mus musculus]	Contains protein domain (PF00435) - Spectrin repeat	52645080, 264691, 264628, 264555
1444	87620478 (2887, 2888)	Novel Protein sim. GBank gl 3874447 emb CAB02772 - (281039) predicted using GeneFinder; cDNA EST EMBL:TO1209 comes from this gene; cDNA EST yk278a11.3 comes from this gene; cDNA EST yk308a9.3 comes from this gene; cDNA EST yk308a9.5 com...	UNCLASSIFIED	264259, 29331822, 29331824, 66714117, 29331828, 264288, 35695917, 33657023, 264635, 60170394
1445	94980470 (2889, 2890)	Novel Protein sim. GBank gl 2959886 emb CAA11022 - (AJ222988) L-peritaxin [Mus musculus]	UNCLASSIFIED	264369
1446	85079268 (2891, 2892)	Novel Protein sim. GBank gl 5081810 gb AAD39464.1 AF13544 - (AF135440) huntingtin yeast partner C [Mus musculus]	Contains protein domain (PF01848) - FF domain	264369
1447	86945392 (2893, 2894)	Novel Protein sim. GBank gl 3980411 AC004561 - putative proline-rich protein [Arabidopsis thaliana]	Contains protein domain (PF00439) - Bromodomain	18108398, 35698286, 22278997, 66714117, 29331828, 265009, 264758, 265018, 264288, 21908766, 21908767, 264692, 264634, 264568
1448	94990477 (2895, 2896)	Novel Protein sim. GBank gl 3980411 AC004561 - putative proline-rich protein [Arabidopsis thaliana]	Contains protein domain (PF00439) - Bromodomain	29331827, 264509, 264909, 265008, 264595, 18108357, 18108385, 264588, 264486

1449	87660859 (2897, 2898)				UNCLASSIFIED	66714117, 284908, 284808, 284591, 284601, 284764, 284832
1450	87456698 (2899, 2900)	Novel Protein sim. GBank gij1707074 (U80450) - M01E11.2 [Caenorhabditis elegans]			UNCLASSIFIED	35696286, 35696052, 265008, 265009, 80170831, 33109954, 284683, 284689, 35696423, 35695855, 56528468
1451	87787970 (2901, 2902)	Novel Protein sim. GBank gij14160304[emb]CAA10600] - (AJ132182) HS1 binding protein 3 [Mus musculus]			UNCLASSIFIED	29331826, 284683, 284693, 263978, 264630
1452	85692899 (2903, 2904)	Novel Protein sim. GBank gij2832900[dbj]BAA24608.1] - (D99340) dipeptidyl peptidase III [Rattus norvegicus]			peptidase	264681, 33657023, 264629
1453	86130434 (2905, 2906)	Novel Protein sim. GBank gij728831[sp]P39188[ALU1_HUMAN - III] ALU SUBFAMILY J WARNING ENTRY IIII			kinase	284510, 264768
1454	11204686 (2907, 2908)					
1455	87787896 (2909, 2910)					
1456	86320218 (2911, 2912)	Novel Protein sim. GBank gij729230[sp]P41004[CUT3_SCHPO - CHROMOSOME SEGREGATION PROTEIN CUT3]			UNCLASSIFIED	264558 29331822, 66714117, 29331825, 284905, 29331830, 265008, 285008, 265009, 265011, 265019, 18108351, 21908768, 33657109, 18108376, 284632, 56182323, 87188518
1457	80076900 (2913, 2914)	Novel Protein sim. GBank gij729230[sp]P41004[CUT3_SCHPO - CHROMOSOME SEGREGATION PROTEIN CUT3]			transport	22278995, 22278998, 22278997, 22278998, 22278999, 29331827, 284107, 265017, 21908765, 21908768, 21908767, 21908769, 29148829, 18108370, 22278000
1458	87800460 (2915, 2916)	Novel Protein sim. GBank gij2246532 (U93872) - ORF 73. contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus]			UNCLASSIFIED	264107, 264568 56182575, 22278998, 60432049, 284259, 29331826, 29331827, 29331828, 285017, 285018, 284107, 264110, 265009, 60432229, 265019, 265020, 263972, 263976, 264635, 22278002, 264568
1459	95360920 (2917, 2918)	Novel Protein sim. GBank gij5524667[gb]AAD44333.1[AF15935] - (AF159356) Munc13-4 protein [Rattus norvegicus]	Contains protein domain (PF00168) - C2 domain		kinase	22278997, 284259, 29331824, 29331826, 29331827, 29331828, 285017, 285018, 264760, 284682, 264448, 264288, 264766, 265021, 284692, 33657023, 33657109, 35695855, 284568
1460	95354602 (2919, 2920)				UNCLASSIFIED	29331822, 264591, 55811957, 264691, 284693, 65274620
1461	94741513 (2921, 2922)	Novel Protein sim. GBank gij1707274 (U80931) - strong similarity to class-III of pyridoxal-phosphate-dependent aminotransferases [Caenorhabditis elegans]	Contains protein domain (PF00202) - Aminotransferases class-III pyridoxal phosphate		gaba	22278997, 29331822, 35696052, 285009, 284758, 285017, 285018, 265019, 284760, 264369, 284687, 21908765, 21908768, 265022, 33657109, 27488261, 284555, 83373044
1462	87732018 (2923, 2924)					
1463	88080605 (2925, 2926)	Novel Protein sim. GBank gij1770466[emb]CAA66912] - (X98259) M-phase phosphoprotein 8 [Homo sapiens]	Contains protein domain (PF00385) - 'chromo' (CHROMALIN Organization MODifier) domain		UNCLASSIFIED	284555, 284558 60432049, 264259, 28146499, 284908, 264907, 284512, 265017, 264763, 284766, 18108370, 18108374, 264636, 18108385, 18108388

1464	87620482 (2927, 2928)	Novel Protein sim. GBank gjl3874447[emb]CAB02772] - (Z81039) predicted using GeneFinder; cDNA EST EMBL:TO1209 comes from this gene; cDNA EST yk278a11.3 comes from this gene; cDNA EST yk278a11.5 comes from this gene; cDNA EST yk308a9.3 comes from this gene; cDNA EST yk308a9.5 com...		UNCLASSIFIED	264559, 22278995, 22278996, 22278997, 22278998, 29331822, 29331824, 29331825, 3569052, 20281100, 264905, 29331830, 264909, 265007, 33657402, 21908754, 265017, 265018, 264882, 264684, 264369, 264286, 264766, 21908765, 21906766, 21908767, 21906769, 35695917, 264691, 33657023, 264692, 35696423, 35695855, 264630, 264631, 264639, 264565, 264488, 22278994, 56994075, 60432049, 264259, 56182181, 60432289, 29331827, 52844045, 264511, 265007, 265008, 264596, 55812038, 55811366, 264600, 264602, 265017, 265018, 264604, 265019, 18108351, 18108354, 56181562, 21908769, 265021, 33657023, 33657182, 55811576, 264557, 18108382, 60432113
1465	87425192 (2929, 2930)	Novel Protein sim. GBank gjl4589598[jb]BAA76821.1] - (AB023194) KIAA0977 protein [Homo sapiens]		glucoamy/lase	264512, 265017, 264689, 264558, 264683, 264636
1466	87605227 (2931, 2932)	Novel Protein sim. GBank gjl2248532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus]		UNCLASSIFIED	
1467	87614328 (2933, 2934)	Novel Protein sim. GBank gjl4507241[ref]NP_003137.1[psRRP - structure specific recognition protein 1]		struct	
1468	95342862 (2935, 2936)	Novel Protein sim. GBank gjl19065586 (U81788) - kinesin-73 [Drosophila melanogaster]		struct	18108394, 18108397, 18108398, 35696052, 29146499, 265007, 265008, 265009, 265010, 265011, 18108354, 18108365, 18108368, 18108374, 18108381, 18108382, 18108384, 18108388
1469	79236174 (2937, 2938)	Novel Protein sim. GBank gjl19065586 (U81788) - kinesin-73 [Drosophila melanogaster]		UNCLASSIFIED	18108394, 18108398, 56182575, 264259, 29331822, 29331824, 29331825, 60432289, 264907, 264909, 265007, 264910, 265009, 264591, 60432229, 60433356, 264595, 60433436, 264758, 33109954, 265010, 265011, 265018, 264760, 264448, 264764, 264288, 264369, 18108357, 264769, 18108358, 21906767, 21906769, 55811957, 265021, 18108361, 264691, 18108362, 18108385, 18108388, 264628, 18108379, 264637, 264557, 18108381, 56182323, 18108382, 83373044, 18108384, 18108388, 87168518, 60432113, 264404, 22279002, 264482, 264567, 264487
1470	94980482 (2939, 2940)	Novel Protein sim. GBank gjl5649170[jb]AAD43131.2[AF159092] - (AF159092) sylt709613 protein [Homo sapiens]		transport	29331822, 29331824, 29331825, 264828, 264603, 264689, 264693, 18108374, 55811576
1471	87826842 (2941, 2942)	Novel Protein sim. GBank gjl3876146[emb]CAB01750] - (Z78542) similar to Mitochondrial carrier proteins; cDNA EST EMBL:TO1651 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00153) - Mitochondrial carrier proteins		

1472	87756616 (2943, 2944)	Novel Protein sim. GBank gi 4680707 gb AAD27743.1 AF13296 - (AF13296) CGI-34 protein [Homo sapiens]		UNCLASSIFIED	264905, 18108351, 21906765, 264486
1473	87791609 (2945, 2946)	Novel Protein sim. GBank gi 3688780 (AF042180) - testis- specific Y-encoded-like protein [Mus musculus]	Contains protein domain (PF00956) - MHC Nucleosome assembly protein (NAP)		18108394, 22278995, 56994075, 22278999, 29331822, 29331824, 86714117, 29331825, 29331826, 35696052, 264906, 264907, 58182435, 265007, 264758, 265018, 265019, 264760, 264764, 264288, 264685, 264688, 264769, 21906769, 55811957, 285021, 264691, 264693, 264629, 55811576, 264634, 264636, 56182323, 22278002, 284566, 284486
1474	85800989 (2947, 2948)	Novel Protein sim. GBank gi 2494890 sp Q92176 CORO_BOVIN - CORONIN-LIKE PROTEIN P57	Contains protein domain (PF00400) - struct WD domain, G-beta repeat		264486, 35695917, 35888286, 264692, 33657023, 264693, 33657109, 35896052, 264508, 264905, 264906, 264907, 264629, 264908, 264909, 35696423, 35695855, 264511, 264910, 264632, 264634, 264635, 264636, 264637, 264556, 264557, 264639, 264758, 60432113, 264604, 264605, 264585, 264586, 264764, 264486, 264685, 264766, 264681, 264682, 264288, 264566
1475	86671935 (2949, 2950)		Contains protein domain (PF00041) - UNCLASSIFIED Fibronectin type III domain		
1476	87548855 (2951, 2952)	Novel Protein sim. GBank gi 475752 ref NP_004664.1 pANGP - angiotensin 3	Contains protein domain (PF00147) - glycoprotein Fibronectin beta and gamma chains, C-terminal globular domain		60424179, 56181886, 29331824, 60424269, 29331826, 35696052, 264508, 264905, 264906, 264907, 264908, 264909, 264512, 265007, 265008, 265009, 264910, 33657402, 264595, 264596, 55812036, 265011, 264601, 264762, 18108351, 264288, 264369, 264685, 264766, 264689, 55811837, 264691, 264692, 264693, 18108370, 60431526, 18108374, 35696423, 264634, 264635, 264636, 60431850, 264555, 264638, 264557, 264639, 18108382, 18108388, 60432113, 22278002
1477	87774279 (2953, 2954)	Novel Protein sim. GBank gi 2498308 sp Q60870 DP1_MOUSE - POLYPOSIS LOCUS PROTEIN 1 HOMOLOG (Tb2 PROTEIN HOMOLOG (GP108))		UNCLASSIFIED	264259, 264107, 264905, 265008, 265010, 265011, 264682, 264288, 265020, 265021, 263974
1478	11754412 (2955, 2956)				264686

1478	91640140 (2957, 2958)	Novel Protein sim. GBank gi 5499741 gb AAD43978.1 AF15298 - (AF152981) chromatin-specific transcription elongation factor FACT 140 kDa subunit [Homo sapiens]		peptidase	56182575, 22278995, 22278996, 22278998, 22278999, 29331822, 29331824, 68714117, 264906, 264907, 56182435, 265006, 60170831, 33657402, 264758, 33109954, 21906754, 265017, 265019, 264448, 264288, 264767, 264687, 52644229, 21906764, 264689, 21906765, 21906768, 21906769, 265020, 265021, 60170815, 264691, 33657023, 33657109, 33657182, 27486281, 27486282, 33657349, 18108370, 60431528, 263976, 55811578, 264556, 264557, 60170394, 87168518, 264404, 22279000, 22279002, 264583, 264482
1480	94312412 (2959, 2960)	Novel Protein sim. GBank gi 3550456 emb CAA06329.1 - (AJ005073) Alix [Mus musculus]		UNCLASSIFIED	18108394, 65274572, 56182575, 22278995, 35696286, 56994075, 22278996, 22278997, 22278998, 22278999, 264091, 264259, 35696052, 28146499, 264103, 264105, 264108, 264907, 52644045, 264112, 265007, 265008, 265009, 60433358, 60433438, 264596, 33109954, 33657084, 52644289, 87168474, 265010, 87168559, 265017, 265018, 265019, 264448, 264682, 264683, 264769, 21906765, 21906768, 21906767, 21906768, 21906769, 265020, 265021, 60170615, 52644150, 33657109, 33657182, 263972, 35695855, 264557, 263981, 83373044, 18108385, 87168518, 264566, 264089, 264486
1481	87021442 (2961, 2962)	Novel Protein sim. GBank gi 4836807 gb AAD30568.1 AF14679 - (AF146793) PFT27 [Mus musculus]		MHC	265008, 265007, 265010, 18108374
1482	85320442 (2963, 2964)	Novel Protein sim. GBank gi 4585372 gb AAD25403.1 AF12292 - (AF122923) Wnt inhibitory factor-1 [Mus musculus]	Contains protein domain (PF00008) - EGF-like domain	UNCLASSIFIED	284908, 284910, 264758
1483	94115503 (2965, 2966)	Novel Protein sim. GBank gi 535428 (U13736) - calmodulin- like protein [Pisum sativum]	Contains protein domain (PF00036) - EF hand	struct	264259, 29331822, 52645080, 29331825, 29331826, 33656970, 29331830, 265007, 55812038, 33109954, 265017, 264286, 21906768, 21906769, 264636, 18108380, 87168518, 22279000
1484	94131544 (2967, 2968)	Novel Protein sim. GBank gi 1911774 bbs 180090 - (S83384) putative Rab5-interacting protein (clone L1-57) [human, HeLa cells, Peptide Partial, 122 aa] [Homo sapiens]		UNCLASSIFIED	264489, 35696286, 264259, 264107, 264809, 265008, 60433356, 33657402, 60433438, 264288, 21906765, 21906768, 29148627, 33657023, 27486262, 18108374, 35696423, 83373044, 60432113
1485	80194441 (2969, 2970)	Novel Protein sim. GBank gi 5360128 gb AAD42883.1 AF15311 - (AF153117) NY-REN 62 antigen [Homo sapiens]	Contains protein domain (PF00225) - Kinesin motor domain	struct	264369, 265020, 18108374

1486	94125068 (2871, 2872)	Novel Protein sim. GBank gi 4589516 dbj BAA76780.1 - (AB023153) KIAA0938 protein [Homo sapiens]	Contains protein domain (PF00089) - Kinase Eukaryotic protein kinase domain	58182575, 22278999, 264908, 284907, 21908754, 87188474, 265017, 265019, 18108351, 264288, 285020, 264568 21908754, 264488
1487	88452711 (2873, 2874)	Novel Protein sim. GBank gi 5018275 emb CAB4443.1 - (AJ132751) xenobiotic/medium-chain fatty acid:CoA ligase form XL-III [Bos taurus]	synthase	
1488	87732028 (2875, 2876)	Novel Protein sim. GBank gi 5712131 gb AADM7378.1 AF12049 - (AF120499) DEM1 protein [Homo sapiens]	Contains protein domain (PF01443) - (gf) Viral (Superfamily 1) RNA helicase	264688, 264789, 284689, 284692, 284683, 284509, 264908, 284907, 18108370, 264908, 264629, 264909, 264510, 285008, 264512, 285007, 265008, 265009, 264555, 264558, 284557, 264558, 264782, 284584, 264882 21808787, 22278999, 285022, 264259
1489	95104277 (2877, 2878)	Novel Protein sim. GBank gi 2487303 sp Q82788 FPRP_RAT - PROSTAGLANDIN F2-ALPHA RECEPTOR REGULATORY PROTEIN PRECURSOR (PROSTAGLANDIN F2-ALPHA RECEPTOR ASSOCIATED PROTEIN)	prostaglandin Immunoglobulin domain	264683, 29331824, 29331825, 29331826, 29331827, 29331828, 264103, 263972, 68712502, 35696423, 35895855, 265007, 265008, 265009, 83373044, 21908754, 58528488, 265017, 264563, 18108351, 284584, 284586, 264369, 284288
1490	87390127 (2879, 2880)		UNCLASSIFIED	58182575, 264259, 29331822, 29331824, 68714117, 29331827, 29331828, 264508, 264905, 68712502, 265007, 265008, 264594, 33857402, 55812038, 87168474, 265018, 18108351, 284369, 264288, 264789, 284689, 21908787, 21908788, 55811957, 80170615, 33857109, 35695855, 264635, 60170394, 56526486, 22279002, 284583 265007, 264448, 18108372, 264558, 58182323
1491	83594305 (2881, 2882)	Novel Protein sim. GBank gi 295671 (L11275) - selected as a weak suppressor of a mutant of the subunit AC40 of DNA dependant RNA polymerase I and III [Saccharomyces cerevisiae]	UNCLASSIFIED	
1492	85805383 (2883, 2884)	Novel Protein sim. GBank gi 1656005 (U71205) - rit [Mus musculus]	Contains protein domain (PF00071) - Ras family oncogene	22278997, 22278998, 29331822, 264907, 68712502

1493	91677215 (2885, 2886)	Novel Protein sim. GBank gi 5689515 dbj BAA03041.1 - (AB029012) KIAA1089 protein [Homo sapiens]		UNCLASSIFIED	264488, 52646365, 65274572, 56182575, 22278994, 35898286, 56994075, 22278999, 60432048, 29331824, 29331828, 35898052, 264508, 264905, 264908, 52644045, 264909, 56182435, 285008, 285009, 60170831, 33657402, 55812038, 265010, 285011, 285017, 265018, 285019, 55811150, 264448, 264882, 264885, 264888, 52644229, 21908765, 21908768, 21908768, 21908769, 265020, 265021, 60170615, 52644150, 33857023, 18108384, 18108385, 33857109, 33857182, 27488261, 27488262, 27488264, 33857349, 27488265, 35695783, 18108370, 264829, 18108374, 52644332, 56182323, 87188518, 22279002, 264584, 264588, 264587
1494	87605265 (2887, 2988)	Novel Protein sim. GBank gi 728832 sp P39189 ALU2_HUMAN - IIII ALU SUBFAMILY SB WARNING ENTRY IIII	Contains protein domain (PF01352) - KRAB box	kinase	264807, 265009, 264768, 18108370, 55811576, 264639, 264565, 264488
1495	87605267 (2889, 2990)	Novel Protein sim. GBank gi 4589588 dbj BAA76816.1 - (AB023189) KIAA0972 protein [Homo sapiens]	Contains protein domain (PF01352) - KRAB box	transcript factor	22278997, 264259, 264908, 264907, 285009, 264594, 33857084, 265017, 264760, 264448, 33857109, 264630, 264634, 56526486, 264583, 264565, 264588, 264488, 264587
1496	87784322 (2891, 2992)	Novel Protein sim. GBank gi 5420387 emb CA846879.1 - (AJ243459) proteophosphoglycan [Leishmania major]		UNCLASSIFIED	35896288, 264908, 265019, 264693
1497	81895428 (2993, 2994)	Novel Protein sim. GBank gi 3874925 emb CAA92591.1 - (Z68296) Similarity to Mouse A-RAF proto-oncogene serine/threonine-protein kinase (SW:KRAA_MOUSE); cDNA EST EMBL:D27610 comes from this gene; cDNA EST EMBL:T01018 comes from this gene; cDNA EST EMBL:D33256 comes from this gene;...	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	kinase	264910, 264758, 265011, 264764, 264288, 264690, 264634, 264635, 56526488
1498	80934938 (2995, 2996)	Novel Protein sim. GBank gi 728836 sp P39193 ALU6_HUMAN - IIII ALU SUBFAMILY SP WARNING ENTRY IIII		oncogene	264488, 65274572, 29331822, 265017, 265018, 21908765, 29148827, 29148829, 18108374, 264637, 264638, 264567
1499	86451569 (2997, 2998)	Novel Protein sim. GBank gi 2570198 (U54556) - microfilarial sheath protein SHP3 [Litomosoides sigmodontis]		glucosylase	263978, 264568
1500	80489386 (2999, 3000)	Novel Protein sim. GBank gi 2078483 (U43200) - antifreeze glycoprotein precursor [Boreogadus saida]		UNCLASSIFIED	22278999, 264789, 18108379
1501	85785297 (3001, 3002)			UNCLASSIFIED	264559
1502	80206141 (3003, 3004)				264508, 264112, 264604, 264684, 52644150, 55811576, 264632, 264556, 264638, 56182323, 264583, 264488
1503	87012701 (3005, 3006)	Novel Protein sim. GBank gi 3900855 (AC004874) - similar to N-acetylgalactosaminyltransferase; similar to Q07537 (P1D1g1171989) [Homo sapiens]	Contains protein domain (PF00535) - Glycosyl transferases	transferase	29331822, 265007, 264369

1504	78840051 (3007, 3008)		Contains protein domain (PF00023) - Ank repeat	UNCLASSIFIED	284693
1505	86102872 (3008, 3010)	Novel Protein sim. GBank gij4753775[embjCAB41870.1] - (AJ132545) protein kinase [Homo sapiens]	Contains protein domain (PF00089) - Eukaryotic protein kinase domain	kinase	29331826, 35696052, 284509, 264908, 284907, 284908, 284909, 284511, 284910, 55812038, 284759, 284763, 284764, 284689, 35695917, 285022, 33657109, 18108374, 284631, 284635, 284638, 284568
1508	94143218 (3011, 3012)	Novel Protein sim. GBank gij1304201[dbjBAA06170] - (D29786) alternatively spliced product [Rattus norvegicus]	Contains protein domain (PF00018) - SH3 domain	glycoprotein	65274572, 56182575, 56994075, 22278997, 22278998, 22278999, 284091, 264092, 60432049, 284259, 52845080, 29331822, 29331827, 284108, 29331830, 284908, 56182435, 284110, 284511, 284512, 55812038, 21908754, 87188559, 284600, 265017, 265018, 284681, 18108354, 284389, 284637, 284689, 21908765, 29146627, 21908768, 21908769, 29148829, 52844150, 33857023, 18108376, 65274791, 56182323, 284558, 284559, 18108385, 87188518, 60432113, 22279000, 284565
1507	83738250 (3013, 3014)	Novel Protein sim. GBank gij5689513[dbjBAA03040.1] - (AB029011) KIAA1088 protein [Homo sapiens]		helicase	284639
1508	11618758 (3015, 3016)	Novel Protein sim. GBank gij5011975[refNP_005875.1]pPAK4 - protein kinase related to S. cerevisiae STE20, effector for Cdc42Hs	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	kinase	264593
1509	87318451 (3017, 3018)	Novel Protein sim. GBank gij113161[spP28614]ACOR - ALCEU - ACETON		UNCLASSIFIED	284259, 29331822, 285007, 18108374, 284556
1510	95362843 (3019, 3020)	Novel Protein sim. GBank gij113161[spP28614]ACOR - ALCEU - ACETON		UNCLASSIFIED	284259, 29331822, 285007, 18108374, 284556
1511	88318073 (3021, 3022)	Novel Protein sim. GBank gij126831[spP39188]ALU1 - HUMAN - IIII ALU SUBFAMILY J WARNING ENTRY IIII		UNCLASSIFIED	284259, 29331822, 29331824, 29331827, 284509, 284907, 284510, 284511, 285007, 284512, 285008, 87188559, 284288, 265022, 33657023, 35695855, 284637, 284638, 284563
1512	85345390 (3023, 3024)	Novel Protein sim. GBank gij4559353[gbjAAD23014.1]AC00658 - (AC006585) putative extragenic suppressor protein [Arabidopsis thaliana]	Contains protein domain (PF01163) - RIO1/ZK632.3/MJ0444 family		52845156, 18108396, 56994075, 60432289, 285008, 60433358, 60433438, 21908754, 87188474, 87188559, 285018, 284762, 284763, 284687, 21906765, 21906769, 27488262, 35695763, 18108374, 35696423, 284555, 18108385, 18108387, 18108388, 87188518, 284482
1513	87436228 (3025, 3026)	Novel Protein sim. GBank gij1330394 (U58761) - C01F1.6 gene product [Caenorhabditis elegans]			35696052, 284905, 284906, 284907, 284908, 284909, 284910, 284591, 284766, 284689, 284692, 284629, 284636

1514	95345392 (3027, 3028)	Novel Protein sim. GBank gi 4559353 gb AA023014.1 AC006585 putative extragenic suppressor protein [Arabidopsis thaliana]	Contains protein domain (PF01163) - RIO1/ZK832.3/MJ0444 family	UNCLASSIFIED	52644507, 52645156, 52646365, 52646842, 65274572, 22278994, 35698288, 56994075, 284259, 52845080, 28331822, 29331825, 35696052, 29331830, 52644045, 56182435, 285008, 60433358, 60433438, 59812038, 21908754, 52646317, 52644298, 87168474, 87168559, 284448, 52644229, 21908785, 21908768, 21908767, 21908768, 35695917, 265020, 52644150, 33857023, 52645129, 33857109, 33857182, 27486261, 27486282, 27486284, 27486285, 35695763, 18108378, 35698423, 35695855, 52644332, 18108385, 18108387, 87168518, 60432113 265020, 264639
1515	79163536 (3028, 3030)	Novel Protein sim. GBank gi 387950 emb CAA87795 - (Z47812) similar to ubiquitin carboxyl-terminal hydrolase; cDNA EST EMBL:D33368 comes from this gene; cDNA EST EMBL:D33965 comes from this gene; cDNA EST EMBL:D33822 comes from this gene; cDNA EST EMBL:D34547 comes from this ge...	ubiquitin	UNCLASSIFIED	
1516	88073539 (3031, 3032)	Novel Protein sim. GBank gi 498015 (L27479) - X123 [Homo sapiens]	UNCLASSIFIED	265008, 56182323, 22279002	
1517	87793325 (3033, 3034)	Novel Protein sim. GBank gi 3415134 (AF082024) - Phyb1 [Pimpinella brachycarpa]		284091, 18108370, 264404	
1518	87350697 (3035, 3036)	Novel Protein sim. GBank gi 728838 sp P39195 ALU8_HUMAN - III ALU SUBFAMILY SX WARNING ENTRY IIII	lm7	58714117, 284508, 284509, 284905, 284510, 284910, 284591, 284595, 284288, 264768, 264769, 18108374, 284638, 284638, 264486 264589, 264489, 60432049, 285009, 33857402, 284598, 21908754, 285019, 284389, 21908785, 21908788, 21908789, 264691, 65274820, 33857182, 27486281, 18108374, 264557, 264639, 87168518, 22279002	
1519	84326689 (3037, 3038)	Novel Protein sim. GBank gi 526268 emb CAB45771.1 - (AL080198) hypothetical protein [Homo sapiens]		UNCLASSIFIED	
1520	87592855 (3039, 3040)	Novel Protein sim. GBank gi 268216 db BAA23712 - (AB007800) HH0452 cDNA clone for KIAA0440 has a 438- bp insertion at position 1711 of the sequence of KIAA0440. [Homo sapiens]	UNCLASSIFIED	18108392, 60432049, 264259, 29331824, 265007, 60433356, 265010, 21908768, 264638	
1521	86970696 (3041, 3042)	Novel Protein sim. GBank gi 5052351 gb AAD38518.1 AF13542 - (AF135421) GDP- mannose pyrophosphorylase B [Homo sapiens]	Contains protein domain (PF00483) - Nucleoside transferase	synthase	18108394, 264259, 66714117, 285011, 264603, 285019, 18108384, 35698423, 264557, 284558, 18108388
1522	78860687 (3043, 3044)	Novel Protein sim. GBank gi 3776567 (AC005388) - Strong similarity to F21B7.33 gi 2809284 from A. thaliana BAC gb AC002560. EST gb N65119 comes from this gene. [Arabidopsis thaliana]	UNCLASSIFIED	29331824, 285018, 285020, 285021	
1523	91005151 (3045, 3046)		UNCLASSIFIED	65274572, 21908768, 284693	
1524	80203723 (3047, 3048)	Novel Protein sim. GBank	UNCLASSIFIED	264112, 21908754, 263974	
1525	87799867 (3049, 3050)	gi 4759040 ref NP_004283.1 pRIN1 - ras inhibitor	UNCLASSIFIED	264683, 284687, 264689, 264690, 264692, 264693	

1526	95105344 (3051, 3052)	Novel Protein sim. GBank gi728850 epP08640 AMYH_YEAST - GLUCOAMYLASE S1/S2 PRECURSOR (GLUCAN 1,4-ALPHA- GLUCOSIDASE) (1,4-ALPHA-D-GLUCAN GLUCOHYDROLASE)			glycoprotein	35698286, 56182181, 60431735, 284595, 55812038, 284605, 284683, 21906785, 55811957, 265020, 65274791, 284555, 264556, 284557, 284558, 284559, 83373044
1527	86262512 (3053, 3054)	Novel Protein sim. GBank gi2792486 (AF041107) - tulip 2 [Rattus norvegicus]				56182575, 284259, 60432049, 29331822, 60432289, 284908, 68712502, 60433436, 87168559, 285017, 284288, 21906768, 21908769, 263977, 55811576, 56182323, 18108381
1528	94130918 (3055, 3056)				UNCLASSIFIED	22278995, 22278997, 284259, 68712502, 264596, 265017, 265019, 284682, 284448, 284683, 284764, 284885, 284686, 21906765, 21906766, 21906767, 21906768, 21906769, 265022, 264693, 83373044, 18108385
1529	94120763 (3057, 3058)	Novel Protein sim. GBank gi4406663 gb AAD20053] - (AF131826) Unknown [Homo sapiens]			UNCLASSIFIED	264488, 263994, 56182575, 22278995, 35698286, 22278997, 284259, 29331822, 60432289, 29331827, 35698052, 284508, 284906, 284907, 284908, 284909, 52844045, 56182435, 284511, 265009, 284910, 60433356, 60433438, 285017, 285018, 264760, 264448, 284764, 284369, 284288, 284768, 18108357, 284768, 52844229, 21906765, 21906766, 21906767, 21906768, 265021, 265022, 52844150, 33657109, 284629, 35695855, 60432113, 22278002, 284563, 284584, 284486, 284587
1530	95012765 (3059, 3060)	Novel Protein sim. GBank gi2828710 (AF043542) - malin cyclophilin [Rattus norvegicus]				264488, 284489, 35698286, 29331825, 35698052, 284508, 284905, 284906, 284907, 284909, 264510, 284511, 264512, 284910, 284592, 284595, 18108351, 264764, 284683, 264684, 284766, 264768, 18108357, 284789, 35695917, 284628, 284629, 18108374, 35695855, 264630, 264631, 284634, 284555, 284636, 284637, 284404, 284563, 284586, 264486
1531	95419351 (3061, 3062)	Novel Protein sim. GBank gi1905874 (U90878) - carboxyl terminal LIM domain protein [Homo sapiens]		Contains protein domain (PF00595) - PDZ domain (Also known as DHR or GLGF).	kinase	56182575, 35696286, 264097, 284259, 29331822, 29331825, 29331826, 29331827, 35698052, 284509, 56182435, 284510, 284511, 285007, 60433358, 55811386, 284681, 284369, 284288, 284766, 284687, 55811957, 35695917, 33657023, 35695783, 55810764, 35696423, 55811578, 283981, 60170394, 56182323, 83373044, 60432113, 284586

1532	85718224 (3063, 3064)	Novel Protein sim. GBank gij3874716[emb CAA91265] - (Z68494) cDNA EST EMBL:D85271 comes from this gene; cDNA EST EMBL:D84845 comes from this gene; cDNA EST EMBL:D84449 comes from this gene; cDNA EST EMBL:D87438 comes from this gene; cDNA EST EMBL:D88087 comes from this gene; cDN... Novel Protein sim. GBank gij1490324[emb CAB01543] - (Z7E141) unknown [Mus musculus]			UNCLASSIFIED	284689
1533	94239830 (3065, 3066)			sinucl		29331824, 29148489, 264907, 264112, 285008, 265011, 265017, 265018, 264782, 18108351, 263967, 20281149, 18108374, 283981, 264588
1534	95343941 (3067, 3068)	Novit Protein sim. GBank gij181286[pir S22697] - extensin - Volvox carter (fragment)			UNCLASSIFIED	264905, 264907, 264766, 264637
1535	80936732 (3069, 3070)					65274572, 22278997, 264259, 60432049, 29331822, 60432289, 29331827, 28148489, 285008, 265008, 60170831, 60433438, 33109954, 87168559, 265018, 18108357, 21908768, 29148828, 265021, 265022, 18108377, 58182323, 60432113, 22278000, 22278002
1536	87602859 (3071, 3072)	Novel Protein sim. GBank gij106024[pir B32891] - finger protein 2, placental - human	Contains protein domain (PF00009) - Zinc finger, C2H2 type	transcript factor		284686, 18108357, 18108394, 21906767, 21908768, 29148828, 35888286, 285020, 265021, 52844150, 264693, 68714117, 29331825, 29331828, 264508, 264905, 20281149, 264909, 18108374, 35898423, 35898555, 265009, 264634, 264638, 264638, 18108385, 56526488, 265017, 265018, 264583, 264762, 18108351, 264448, 264369, 264768
1537	95354556 (3073, 3074)	Novel Protein sim. GBank gij3876332[emb CAB02096] - (Z78754) cDNA EST EMBL:T01054 comes from this gene; cDNA EST EMBL:D73600 comes from this gene; cDNA EST yk426f12.5 comes from this gene; cDNA EST yk342f10.5 comes from this gene; cDNA EST yk475c5.5 comes from this gene; cDNA ES...				65274572, 58182575, 60432049, 284259, 29331828, 285008, 285007, 60433356, 60433438, 264601, 18108351, 264448, 264369, 264288, 33857023, 65274620, 33857109, 60432113
1538	85724828 (3075, 3076)	Novel Protein sim. GBank gij403440 (M81787) - [Gallus domesticus skeletal muscle mRNA, partial cds.] gene product [Gallus gallus]	Contains protein domain (PF00009) - Eukaryotic protein kinase domain	kinase		18108394, 18108397, 264909, 265008, 285009, 265010, 18108351, 284638, 18108382, 18108385, 18108388

1539	95337828 (3077, 3078)	Novel Protein sim. GBank gij3218411lembi(CAA19575.1) - (AL023859) SPBC19C7.07c, putative tRNA splicing endonuclease gamma subunit, len:284aa, similar eg. to YAR008W, YAH8_YEAST, P39707, YAR008W, tRNA splicing endonuclease gamma subunit, (275aa), fasta scores, opt:269, E(1):8.4e-2...		nuclease	22278994, 22278996, 35896286, 58994075, 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 60424289, 29331825, 60432269, 29331826, 29331827, 29331828, 35896052, 33858970, 58182435, 285009, 33857402, 60433356, 60433438, 55812038, 21906754, 87168559, 285017, 285018, 265019, 18108351, 284288, 52644229, 18108359, 21908784, 21908787, 21908788, 35895917, 265020, 265021, 52644150, 33857023, 33857109, 27486281, 18108370, 18108376, 35896423, 55811578, 65274791, 284558, 58182323, 80170394, 83373044, 87168518, 60432113, 22279000, 22279002, 264568
1540	95352858 (3079, 3080)	Novel Protein sim. GBank gij5052634jpbjAAD38847.1jAF14567 - (AF145672) BcDNA, GH12174 [Drosophila melanogaster]		UNCLASSIFIED	264369, 264691, 263978
1541	95317848 (3081, 3082)	Novel Protein sim. GBank gij5052349jpbjAAD38515.1jAF13501 - (AF135016) protein phosphatase 2A 48 kDa regulatory subunit [Homo sapiens]		phosphatase	264488, 264489, 22278998, 264259, 29331822, 35898052, 264508, 264509, 264905, 264907, 264908, 264511, 264512, 264910, 264592, 264781, 264762, 264448, 264764, 264288, 264687, 21808788, 55811957, 35895917, 265020, 264691, 33857023, 264692, 33857109, 264628, 18108374, 264632, 264634, 264635, 264639, 18108385, 264563, 264564, 264585, 264588, 264488
1542	90837549 (3083, 3084)	Novel Protein sim. GBank gij505702jpbjAAD41779.1jAF12686 - (AF126867) calpain-like protease [Mus musculus]		calthepsin	18108392, 18108394, 65274572, 29331822, 264508, 265007, 265008, 265009, 265011, 264682, 18108354, 18108355, 52644150, 18108368, 264636, 18108381, 18108382
1543	84348768 (3085, 3086)	Novel Protein sim. GBank gij728832jpbjP39189jALU2_HUMAN - III ALU SUBFAMILY SB WARNING ENTRY IIII		nuclease	264108, 33109954, 265019, 264683, 35895917, 264690, 264692, 33857109
1544	87757285 (3087, 3088)	Novel Protein sim. GBank gij3493162 (AF084259) - bromodomain-containing protein BP75 [Mus musculus]	Contains protein domain (PF00439) - Bromodomain	nud_rept	35898286, 56994075, 22278999, 35896052, 60433356, 60433438, 265011, 264683, 33857109, 35898423, 264631, 87168518, 22279000
1545	85757973 (3089, 3090)	Novel Protein sim. GBank gij1086591 (U41007) - similar to S. cerevisiae nuclear protein SNF2 (SP-P22082) in a region of gly-arg repeats [Caenorhabditis elegans]		UNCLASSIFIED	264112, 264692, 264693, 55811578
1546	78476589 (3091, 3092)			UNCLASSIFIED	264805, 264686
1547	88995984 (3093, 3094)	Novel Protein sim. GBank gij2681132 (AF035683) - p21 [Mus musculus]		UNCLASSIFIED	264259, 29331822, 66714117, 265007, 55811388, 265010, 264600, 265017, 265019, 264288, 264788, 265020, 265022, 55811578, 18108380, 264563

1548	94233065 (3095, 3098)	Novel Protein sim. GBank gi 3043692 dbj BAA25510 - (AB011156) KIAA0584 protein [Homo sapiens]		UNCLASSIFIED	29331824, 60431528, 284639, 56182323
1549	95330048 (3097, 3098)	Novel Protein sim. GBank gi 5689519 dbj BAA83043.1 - (AB029014) KIAA1091 protein [Homo sapiens]		eph	60424179, 22278995, 35696286, 22278998, 22278999, 264092, 264094, 29331822, 56182181, 29331824, 35696052, 264905, 284908, 264908, 264909, 265008, 264511, 285008, 60431735, 60433356, 21908754, 55811388, 87168559, 265017, 285018, 285019, 55811150, 264682, 264288, 284389, 56181582, 284769, 21908765, 21908780, 21908789, 55811957, 265020, 264691, 33657109, 60431528, 35696423, 35695855, 56526488, 60432113, 22279002, 284583, 264588
1550	95201907 (3099, 3100)	Novel Protein sim. GBank gi 544483 sp P33350 GUSB_BOVIN - POSSIBLE GUSTATORY RECEPTOR TYPE B (PPR1 PROTEIN)	Contains protein domain (PF00001) - 7 transmembrane receptor (rhodopsin family)	lm7	55274572, 60432289, 265008, 284910, 285011, 265017, 265019, 264768, 56182323
1551	88077111 (3101, 3102)	Novel Protein sim. GBank gi 4758588 ref NP_004798.1 pHS6S - heparan-sulfate 6-sulfotransferase		UNCLASSIFIED	22278999, 29331822, 264508, 264509, 264908, 264907, 264909, 265007, 264512, 284910, 21908754, 285018, 285018, 284681, 284784, 264766, 264688, 264769, 21906769, 284692, 35695763, 284635, 284555, 284559, 284557, 264638, 264558, 264563
1552	87617114 (3103, 3104)			UNCLASSIFIED	284259, 29331828, 66712502, 284784, 284288, 264688, 33657109, 264558
1553	94725512 (3105, 3106)	Novel Protein sim. GBank gi 4589570 dbj BAA76807.1 - (AB023180) KIAA0963 protein [Homo sapiens]	Contains protein domain (PF00304) - Gamma-thionins family	dehydrogenase	56182375, 35696286, 29146499, 284509, 264907, 264908, 264909, 56182435, 285006, 285008, 265009, 264910, 264757, 284758, 265017, 55811150, 18108351, 264784, 56181582, 35695917, 264693, 33657109, 18108374, 35696423, 65274791, 35695855, 284635, 264555, 56182323, 18108382, 83373044, 22279000
1554	94233069 (3107, 3108)	Novel Protein sim. GBank gi 3043692 dbj BAA25510 - (AB011156) KIAA0584 protein [Homo sapiens]	Contains protein domain (PF00446) - Gonadotropin-releasing hormones	transferase	35696286, 22278997, 264259, 29331822, 29331824, 29331825, 29331828, 265007, 265009, 60432229, 33657402, 55812038, 285011, 285019, 284681, 284369, 284686, 284767, 284768, 21908765, 21906769, 35695917, 284693, 18108370, 60431528, 55811576, 264631, 60170394, 56182323, 83373044, 18108385, 22279000, 22279002

1555	87332970 (3109, 3110)	Novel Protein sim. GBank g 2257495 dbj BAA21392 - (AB004534) p 015 [Schizosaccharomyces pombe]	(Contains protein domain (PF00400) - UNCLASSIFIED WD domain, G-beta repeat	284259, 29331826, 35696052, 284508, 284905, 284908, 284907, 284908, 52844045, 284909, 284910, 60432229, 60433358, 55812038, 284758, 284759, 33857084, 285011, 87168559, 284601, 285018, 285019, 284763, 284764, 284288, 284768, 284768, 21908785, 35695917, 285022, 284691, 33657023, 35698423, 35695855, 284635, 284555, 284638, 284638, 284639, 18108385, 56528486
1556	91220268 (3111, 3112)		UNCLASSIFIED	83373044, 284758, 285022, 284600, 35696052, 284630, 35698423, 285018, 284632, 284682, 28331822, 285020, 285011
1557	87640609 (3113, 3114)	Novel Protein sim. GBank g 3329611 (AF078783) - contains similarity to C3HC4-type zinc fingers (Pfam; z C3HC4.hmm, score: 34.08); most similar to drosophila goliath protein (SW: Q06003) [Caenorhabditis elegans]	(Contains protein domain (PF00097) - UNCLASSIFIED Zinc finger, C3HC4 type (RING finger)	60432289, 284509, 284908, 284907, 284908, 284909, 284910, 284758, 55811388, 284761, 284762, 284768, 284769, 284690, 263978, 284634, 284635, 284639, 284584, 284488
1558	94840376 (3115, 3116)	Novel Protein sim. GBank g 5360105 gb AAD42871.1 AF155105 putative zinc finger protein NY-REN-34 antigen [Homo sapiens]		22278994, 22278998, 22278997, 22278998, 22278999, 60432049, 284259, 29331824, 29331825, 29331826, 29331827, 284908, 284909, 60433358, 21908754, 285017, 285018, 284448, 21908767, 285021, 285022, 33657023, 33657108, 18108370, 55811576, 83373044, 87188518, 22278900, 22278902
1559	88224865 (3117, 3118)	Novel Protein sim. GBank g 112908 sp P02750 A2GL_HUMAN - LEUCINE-RICH ALPHA-2-GLYCOPROTEIN (LRG)	(Contains protein domain (PF00560) - glycoprotein Leucine Rich Repeat	284259, 29331822, 60432289, 35696052, 284107, 284110, 21908754, 33109954, 87188559, 284760, 284763, 21908764, 21908785, 21908768, 285021, 284690, 35695855
1560	84580875 (3119, 3120)	Novel Protein sim. GBank g 3880146 emb CAA92704 - (Z68319) Similarity to Human hnRNP F protein (PIR Acc. No. S43484); cDNA EST EMBL:D34218 comes from this gene; cDNA EST EMBL:D37248 comes from this gene; cDNA EST EMBL:D71817 comes from this gene; cDNA EST EMBL:D74531 comes from...	UNCLASSIFIED	284908, 284803, 284638
1561	86609159 (3121, 3122)		UNCLASSIFIED	284510, 284594
1562	83359682 (3123, 3124)		UNCLASSIFIED	283987
1563	85508894 (3125, 3126)		UNCLASSIFIED	284910, 284784, 284768
1564	87766371 (3127, 3128)	Novel Protein sim. GBank g 1168287 sp P45953 ACD_V_RAT - ACYL-COA DEHYDROGENASE, VERY-LONG-CHAIN SPECIFIC PRECURSOR (VLCAD)	(Contains protein domain (PF00441) - dehydrogenase Acyl-CoA dehydrogenase	18108394, 35698286, 284259, 29331822, 60432289, 35696052, 29331826, 284508, 68712502, 284908, 56182435, 285007, 284910, 80170831, 21908754, 285011, 285017, 285018, 285019, 18108351, 284448, 284288, 284886, 21908765, 21908768, 285021, 60170815, 284692, 35698423, 35695855, 284557, 58182323, 60432113, 22278902, 284482

1565	67783381 (3129, 3130)	Novel Protein sim. GBank glj128726[sp]p05307[POI_BOVIN - PROTEIN DISULFIDE ISOMERASE PRECURSOR (PDI) (PROLYL 4-HYDROXYLASE BETA SUBUNIT) (CELLULAR THYROID HORMONE BINDING PROTEIN) (P55)]		isomerase	264488, 264689, 18108398, 55811957, 264534, 264259, 264508, 264905, 264509, 264908, 18108372, 264510, 264511, 264512, 265008, 264630, 265009, 264910, 264635, 264638, 264591, 264555, 264592, 264637, 264593, 264594, 264595, 264596, 265011, 264603, 22279002, 18108351, 264782, 264565, 264567
1566	87424749 (3131, 3132)	Novel Protein sim. GBank glj3800445[emb][CAA20329] - (AL031268) VM106R.1 [Caenorhabditis elegans]		inf	22278986, 22278989, 264259, 29331822, 29331824, 60432288, 29331827, 68712502, 264908, 265008, 18108351, 52644229, 21908765, 21908767, 21908768, 21908769, 33857109, 264555, 264639, 264482
1567	84999006 (3133, 3134)	Novel Protein sim. GBank glj4929899[gb]AAD34110.1[AF15187 - (AF151873) CGI-115 protein [Homo sapiens]		UNCLASSIFIED	56182575, 21906769, 264692
1568	87648761 (3135, 3136)	Novel Protein sim. GBank glj4827063[re]NP_005072.1[pZNF1 - zinc finger protein 142 (clone pHZ-49)]	Contains protein domain (PF000096) - Zinc finger, C2H2 type	transcriptfactor	29331827, 29331830, 264511, 265009, 264758, 21908767, 21908768, 264691, 264893, 22279000, 22279002
1569	90936668 (3137, 3138)	Novel Protein sim. GBank glj5689451[dbj]BAA83009.1 - (AB028880) KIAA1057 protein [Homo sapiens]	Contains protein domain (PF00443) - Ubiquitin carboxyl-terminal hydrolase family 2	ubiquitin	65274572, 29331822, 29331824, 29331828, 264905, 56182435, 265007, 265019, 264784, 21908765, 21908769, 55811957, 60170815, 52644150, 264692, 33857023, 33857109, 18108377, 264563, 264567
1570	86843981 (3139, 3140)	Novel Protein sim. GBank glj1255430 (U53155) - No definition line found [Caenorhabditis elegans]		UNCLASSIFIED	264595, 264682, 265021
1571	91210340 (3141, 3142)	Novel Protein sim. GBank glj4507731[re]NP_001061.1[pTUBG - tubulin, gamma polypeptide]	Contains protein domain (PF000091) - Tubulin/FtsZ family	tubulin	22278986, 35696288, 22278997, 264091, 264259, 29331824, 29331825, 29331827, 35696052, 264508, 264905, 56182435, 264510, 265007, 264758, 265011, 18108351, 264448, 264288, 264369, 21908765, 21908767, 21908768, 21908769, 35695917, 265020, 265021, 33657023, 264693, 18108370, 18108377, 35696423, 35695855, 264634, 264555, 264558, 18108384
1572	81201664 (3143, 3144)			UNCLASSIFIED	35696052, 264905, 264906, 264908, 264910, 264758, 264768, 35695917, 264637
1573	80207066 (3145, 3146)			UNCLASSIFIED	263972

1574	94216142 (3147, 3148)	Novel Protein sim. GBank gi 4758334 ref NP_004256.1 pFADS - delta-6 fatty acid desaturase	Contains protein domain (PF00173) - Heme-binding domain in cytochrome b5 and oxidoreductases	Cytochrome	18108394, 264887, 18108397, 18108398, 22278996, 22278997, 22278999, 264259, 29331825, 29331827, 29146498, 29146499, 264107, 264907, 264909, 52644045, 264511, 265008, 264910, 265009, 264591, 21906754, 265011, 265019, 18108351, 264682, 264763, 264764, 18108354, 264369, 264288, 264885, 264768, 264688, 264768, 264688, 21908765, 21908768, 21908767, 21908768, 21908769, 29148629, 264680, 264691, 264693, 20281089, 18108370, 18108374, 18108378, 35695855, 264634, 18108384, 18108385, 22278002, 264563, 264568
1575	95340019 (3149, 3150)	Novel Protein sim. GBank gi 3891810 emb CAA94856 - (Z70783) similar to EF-hand calcium binding protein; cDNA EST EMBL: C08700 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00038) - EF hand	phosphatase	56994075, 264259, 29331822, 29331824, 29331825, 60432289, 29331828, 264909, 265006, 265008, 265009, 265010, 87168559, 55811150, 264448, 18108354, 264369, 264288, 18108357, 55811857, 265020, 265021, 60170815, 264691, 33657023, 33657109, 60431528, 65274781, 35695855, 18108385, 60432113, 22278002, 264482
1576	95314019 (3151, 3157)	Novel Protein sim. GBank gi 2773195 (AF039711) - contains similarity to Physcomitrella patens glyceraldehyde 3-phosphate dehydrogenase (GB:X7238.1) [Caenorhabditis elegans]		UNCLASSIFIED	264569, 264092, 264094, 264095, 264259, 264508, 264905, 264509, 264907, 264909, 264510, 264511, 265006, 264910, 21906754, 265010, 265011, 87168559, 264761, 264762, 264288, 264768, 264769, 264691, 264693, 35695855, 264632, 264634, 264635, 264638, 83373044, 264486
1577	87613800 (3153, 3154)	Novel Protein sim. GBank gi 2499130 p P70315 WASP_MOUSE - WISKOTT- ALDRICH SYNDROME PROTEIN HOMOLOG (WASP)		UNCLASSIFIED	264511, 265011, 264681, 264369, 264686, 264689, 264629, 264555, 264558, 264559
1578	87123138 (3155, 3156)				264259, 29331826, 265017, 264689, 264693, 60432113
1579	88085141 (3157, 3158)	Novel Protein sim. GBank gi 2978255 dbj BAA25190 - (AB007407) myeloid zinc finger protein-2 [Mus musculus]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	UNCLASSIFIED	35696286, 264908, 264909, 60433438, 55811386, 264369, 264685, 33657023, 264555, 264558, 264557, 87168518
1580	87255702 (3159, 3160)	Novel Protein sim. GBank gi 4324682 gb AAD16986 - (AFJ08674) late gestation lung protein 1 [Rattus norvegicus]	Contains protein domain (PF00188) - SCP-like extracellular protein	glycoprotein	22278999, 35696052, 29331830, 52644045, 55812038, 87168474, 265018, 264448, 265022, 264638, 58526486, 22279000
1581	95087431 (3161, 3162)	Novel Protein sim. GBank gi 2088838 (AF003386) - F59E12.4 gene product [Caenorhabditis elegans]			22278995, 29331822, 29331824, 29331826, 56182435, 264595, 55812038, 87168559, 265017, 264288, 21908764, 55811857, 35695917, 264692, 55811576, 264637, 56182323, 264559, 83373044, 60432113

1582	95356052 (3163, 3164)	Novel Protein sim. GBank gij5420307[embjCAB46879.1] - (AJ243459) proteophosphoglycan [Leishmania m.]		phosphatase	264259, 60432289, 29331827, 264509, 264905, 264906, 264907, 264909, 264910, 264782, 264288, 264788, 264769, 264832, 264555, 264639, 56526488, 22279000
1583	87622715 (3165, 3166)	Novel Protein sim. GBank gij5578958[embjCAB51351.1] - (AL050306) dJ47587.2 (novel protein) [Homo sapiens]		UNCLASSIFIED	60170831, 33657402, 264682, 21806766, 35895855, 264563
1584	95337722 (3167, 3168)	Novel Protein sim. GBank gij5531815[gblAAD44482.1] - (AF078850) steroid dehydrogenase homolog [Homo sapiens]	Contains protein domain (PF00108) - short chain dehydrogenase	dehydrogenase	60424179, 52846842, 65274572, 56182575, 22278995, 35896286, 22278996, 22278998, 22278999, 29331822, 56182181, 60424288, 60432289, 29331827, 29331828, 35696052, 28146498, 68712502, 29331830, 52644045, 56182435, 264510, 264512, 265008, 60433356, 33657402, 60433438, 55812038, 21806754, 5581386, 52644298, 87188474, 87188559, 265018, 265019, 264448, 264369, 264288, 18108358, 21808765, 21808767, 21808788, 21808769, 35895917, 265020, 265021, 265022, 52644150, 33857023, 33857109, 18108374, 55810784, 55811578, 35896423, 65274781, 35895855, 56182323, 83373044, 18108387, 87188518, 60432113, 22279002
1585	87626117 (3169, 3170)	Novel Protein sim. GBank gij4240132[dbjBAA74846.1] - (AB020630) KIAA0823 protein [Homo sapiens]	Contains protein domain (PF00023) - Ank repeat	phosphatase	35695286, 22278998, 264259, 29331822, 29331824, 29331825, 264905, 265008, 265007, 265008, 60433356, 33109954, 87188474, 265011, 265017, 284604, 264369, 264288, 264685, 264769, 18108359, 21808765, 18108364, 18108370, 264829, 263972, 18108383, 18108388, 264482, 264584
1586	88067081 (3171, 3172)	Novel Protein sim. GBank gij3786494 (AF098993) - No definition line found [Caenorhabditis elegans]			265017, 265018, 264689, 33657023, 263978, 264638, 264563
1587	87617126 (3173, 3174)	Novel Protein sim. GBank gij3253159 (AF005355) - translation initiation factor eIF2C [Oryctolagus cuniculus]		UNCLASSIFIED	264907, 264908, 264511, 264910, 264591, 264594, 264629, 264631, 264563, 264483, 264567
1588	87802538 (3175, 3176)	Novel Protein sim. GBank gij1077573[pirjSS2680 - probable ribosomal protein L34, mitochondrial - yeast (Saccharomyces cerevisiae)]	Contains protein domain (PF00468) - Ribosomal protein L34	UNCLASSIFIED	264259, 29331828, 264905, 265008, 264758, 21806754, 264761, 264762, 21806765, 21808769, 60170815, 52844150, 33657109, 35895855, 56182323, 18108385
1589	80980853 (3177, 3178)	Novel Protein sim. GBank gij2137756[pirjI48746 - semaphorin C - mouse (fragment)]		UNCLASSIFIED	65274572, 264490, 29331822, 66714117, 29331827, 29331828, 56182435, 265008, 60170831, 264595, 264758, 264598, 265011, 264686, 21808766, 21808768, 55811957, 27486265, 264639, 18108385, 56526489, 60432113

1590	85318825 (3179, 3180)			UNCLASSIFIED	264489, 22278998, 264259, 29331824, 29331825, 29331826, 29331827, 285006, 60433356, 21906754, 265017, 265018, 265019, 264448, 264765, 264288, 52844228, 21906765, 21906767, 21906768, 21906769, 265021, 264692, 27486285, 35695763, 56526486, 60432113, 22279000, 22279002, 264564
1591	86877160 (3181, 3182)	Novel Protein sim. GBank		MHC	264259, 264905, 29331830, 264595, 265017, 264448, 264288, 264690, 264629, 87168518
1592	87892533 (3183, 3184)	glj4557749[ref]NP_000237.1pMHC2 - MHC class II transactivator			
1593	94891661 (3185, 3186)			UNCLASSIFIED	65274572, 60432049, 264509, 60433356, 21906754, 21906767, 21906768, 18108370, 35696423, 22279000, 264585, 264567
1594	8773752 (3187, 3188)	Novel Protein sim. GBank glj3877072[emb]CAA87080] - (Z46937) similarity with ribosomal protein L21 [Caenorhabditis elegans]		UNCLASSIFIED	264488, 29331827, 264905, 264906, 264807, 264908, 264909, 264910, 264592, 264593, 264757, 264602, 264604, 264760, 264681, 264288, 264766, 264768, 29148629, 35695917, 264692, 264628, 264629, 264630, 264632, 264634, 264635, 264636, 264639, 264563, 264564, 264566
1595	79918425 (3189, 3190)	Novel Protein sim. GBank glj3152703 (AF065389) - tetraspan NET-4 [Homo sapiens]	Contains protein domain (PF00335) - 4 transmembrane segments Integral membrane proteins	UNCLASSIFIED	29331826, 264908, 55811957
1596	78933828 (3191, 3192)			UNCLASSIFIED	29148498, 264758, 263967
1597	86971657 (3193, 3194)	Novel Protein sim. GBank glj5257114[gb]AAD41244.1JAF09448 - (AF094480) cholesterol 24-hydroxylase [Homo sapiens]	Contains protein domain (PF00067) - Cytochrome P450		264092, 29331824, 264508, 264682, 264369, 264686, 264630, 264583
1598	87862939 (3195, 3196)				264259, 264634
1599	87649829 (3197, 3198)	Novel Protein sim. GBank glj4508797[ref]NP_000324.1pSCA7 - spinocerebellar ataxia 7 (olivopontocerebellar atrophy with retinal degeneration)	Contains protein domain (PF01581) - FMRFamide related peptide family	UNCLASSIFIED	52845080, 29331824, 29331826, 264511, 265009, 265011, 264605, 264448, 264764, 265020, 264692, 264693, 18108370, 264635, 18108385
1600	80056002 (3199, 3200)				29331826, 264603, 264691, 264563
1601	15023246 (3201, 3202)				264635
1602	86926987 (3203, 3204)	Novel Protein sim. GBank glj5305704[gb]AAD41780.1JAF12853 - (AF128535) cytoplasmic phosphoprotein PACSIN2 [Mus musculus]	Contains protein domain (PF00016) - SH3 domain	UNCLASSIFIED - struct	29148499, 264112, 264762, 18108351, 29148627, 263974
1603	80502072 (3205, 3206)	Novel Protein sim. GBank glj283920[pir]S27939 - lensin - chicken		collagen	264490, 29331824, 264807, 264909, 264511, 265008, 264592, 265010, 265011, 264762, 264764, 264369, 264288, 264687, 264769, 264693, 264628, 264634, 264636, 264555, 264556, 264638, 264557, 264558, 264559, 18108385
1604	80221813 (3207, 3208)	Novel Protein sim. GBank glj4768831[gb]AAD29633.1JAF11682 - (AF116827) unknown [Homo sapiens]		ATPase associated	18108385, 263977

1605	91221129 (3209, 3210)				struct	264905, 264509, 264908, 264907, 264908, 264909, 264904, 264768, 264768, 264892, 264693, 33657109, 264629, 35895855, 264635, 264638, 264637
1606	94312703 (3211, 3212)	Novel Protein sim. GBank gll4505313[re]NP_003794.1pMYOM - UNKNOWN	Contains protein domain (PF00047) - Immunoglobulin domain	struct	22278998, 22278999, 264259, 33657402, 265017, 18108351, 264448, 21808767, 21908769, 52844150, 264691, 87168518	
1607	10871805 (3213, 3214)	Novel Protein sim. GBank gll5174473[re]NP_005888.1[p]PPI - Intracisternal A particle-promoted polypeptide		transcriptfactor	264689	
1608	80428900 (3215, 3216)	Novel Protein sim. GBank gll2224629[db]BAA20802 - (AB002342) KIAA0344 [Homo sapiens]		UNCLASSIFIED	264094, 264908, 264907, 264909, 264910, 264591, 264603, 264768, 264693, 264634, 264635, 264637, 264639	
1609	94311572 (3217, 3218)	Novel Protein sim. GBank gll4884073[emb]CAB43213.1 - (AL049934) hypothetical protein [Homo sapiens]			52844507, 52845158, 52846365, 52846842, 56182575, 22278994, 56994075, 35898288, 22278997, 22278998, 22278999, 264259, 52845080, 29147820, 29331828, 35898052, 33656970, 264508, 264509, 264907, 52844045, 56182435, 264510, 264511, 264512, 33657402, 21908754, 52846317, 33109954, 52844298, 87168474, 265017, 265018, 265019, 18108351, 264448, 264288, 264769, 52844229, 21908765, 21908788, 21908787, 21908768, 21908769, 55811957, 35895917, 265020, 265021, 265022, 52844150, 33657023, 33657109, 52845129, 27486261, 27486262, 35895783, 264628, 18108370, 18108376, 35898423, 264638, 52844332, 18108387, 87168518, 22278000, 264583, 264488	
1610	85468200 (3219, 3220)	Novel Protein sim. GBank gll283920[pir]S27839 - tensin - chicken		UNCLASSIFIED	264593, 264757, 55812038, 265018, 265020, 264691, 264692, 264693, 264631, 264634, 264635, 264555, 22278000, 264564	
1611	94122843 (3221, 3222)	Novel Protein sim. GBank gll107284[pir]A35415 - peroxidase (EC 1.11.1.7), thyrold (Grave's disease) - human (fragment)	Contains protein domain (PF00008) - EGF-like domain	peroxidase	35898288, 21908765, 264691, 35898423	
1612	85746031 (3223, 3224)	Novel Protein sim. GBank gll3874846[emb]CAA94337 - (Z70307) Similarity to B. subtilis tetracycline resistance protein (SW:TCR2_BACSU); cDNA EST EMBL:C09951 comes from this gene; cDNA EST EMBL:C08265 comes from this gene [Caenorhabditis elegans]		UNCLASSIFIED	264488, 264509, 18108370, 18108387, 264488	
1613	82247354 (3225, 3226)			UNCLASSIFIED	264759	

1614	91228634 (3227, 3228)	Novel Protein sim. GBank gi 4680673 gb AAD27726.1 AF132951) CGI-17 protein [Homo sapiens]	Contains protein domain (PF01605) - eRF1-like proteins	UNCLASSIFIED	22278995, 22278996, 22278997, 22278998, 22278999, 284259, 29331822, 284908, 284512, 265009, 265011, 265017, 265018, 265019, 18108351, 284683, 284288, 284768, 21908767, 21908768, 21908769, 35695917, 265021, 265022, 35696423, 35695855, 60170394, 58182323, 83373044, 264568
1615	86121909 (3228, 3230)	Novel Protein sim. GBank gi 5688485 db BAA83026.1 - (AB028987) KIAA1074 protein [Homo sapiens]	Contains protein domain (PF00023) - Ank repeat	homeobox	22278996, 35696286, 22278997, 29331822, 35696052, 29331828, 284508, 284908, 284909, 56182435, 264511, 265017, 265019, 284768, 284787, 284768, 265020, 284691, 284628, 284632, 284635, 284555, 284558, 58182323, 284558, 22279002
1616	94311819 (3231, 3232)	Novel Protein sim. GBank gi 3876260 emb CAB01688 - (Z78416) cDNA EST EMBL:D71020 comes from this gene; cDNA EST EMBL:D73593 comes from this gene; cDNA EST EMBL:C07849 comes from this gene; cDNA EST EMBL:C09081 comes from this gene; cDNA EST yk399f2.3 comes from this gene; cDNA ...	UNCLASSIFIED	UNCLASSIFIED	284488, 52844507, 52843158, 52846365, 52846842, 22278994, 22278995, 35696286, 22278996, 22278997, 22278999, 52845080, 29331822, 29331824, 29331825, 29331827, 29331828, 35696052, 33656970, 284905, 284909, 284584, 52846317, 21906754, 33657094, 52844286, 87168474, 87168559, 265017, 265018, 285019, 284681, 284448, 284884, 52844229, 21906784, 284689, 21906765, 21906768, 21906769, 35695917, 265020, 265021, 52844150, 33657023, 52845128, 33657109, 33657182, 27486281, 27486282, 33657349, 27486265, 35695783, 18108376, 35698423, 35695855, 264557, 52844332, 284558, 18108385, 87168518
1617	88090742 (3233, 3234)	Novel Protein sim. GBank gi 468053 sp P34679 O41_CAEEL - HYPOTHETICAL 68.7 KD PROTEIN ZK757.1 IN CHROMOSOME III	Contains protein domain (PF01520) - DHC zinc finger domain	peptidase	35696052, 284605, 284509, 284907, 284908, 284310, 284511, 264784, 284768, 284768, 284689, 264693, 18108374, 284635, 284636, 284638
1618	86272860 (3235, 3236)	Novel Protein sim. GBank gi 4240231 db BAA74894.1 - (AB020678) KIAA0871 protein [Homo sapiens]		stuct	35696286, 22278999, 264092, 29331824, 29331825, 35696052, 33657084, 21906765, 27486284
1619	95354580 (3237, 3238)	Novel Protein sim. GBank gi 5031763 ref NP_005515.1 pHRY - hairy (Drosophila)- homolog	Contains protein domain (PF00010) - Helix-loop-helix DNA-binding domain	transcription factor	52846842, 65274572, 22278999, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 35696052, 58182435, 265007, 265008, 284910, 60170831, 60432229, 60433358, 60433438, 265019, 284448, 284288, 284688, 21906768, 265021, 60170815, 33657023, 65274620, 33657109, 18108374, 18108376, 35696423, 35695855, 58182323, 58526486
1620	87344655 (3239, 3240)	Novel Protein sim. GBank gi 1351047 sp P45843 SCRT_DROME - SCARLET PROTEIN		UNCLASSIFIED	284684

1621	87076708 (3241, 3242)	Novel Protein sim. GBank gl 3822553 (AF098788) - nuclear calmodulin-binding protein [Gallus gallus]	Contains protein domain (PF00622) - SPRY domain	UNCLASSIFIED	264910 18108392, 65274572, 18108398, 22278996, 22278997, 22278999, 29146498, 29148499, 264905, 264908, 264909, 264828, 52844045, 264592, 60433358, 21906754, 264602, 265017, 264389, 21906768, 55811957, 265021, 60170615, 264635, 264557, 60170394, 83373044, 18108385, 22279000, 22279002, 264568
1622	84741739 (3243, 3244)				18108398, 264259, 264909, 56182435, 87168474, 264448, 21906768, 35895917, 264691, 87168518, 264583
1623	8779106 (3245, 3246)	Novel Protein sim. GBank gl 731088 sp P40389 UV22_SCHPO - UV-INDUCED PROTEIN UV22		ribosomalprot	264758
1624	87338178 (3247, 3248)	Novel Protein sim. GBank gl 387566 emb CAB05478 - (Z83104) cDNA EST EMBL: T00015 comes from this gene; cDNA EST EMBL: D33685 comes from this gene; cDNA EST EMBL: D36540 comes from this gene; cDNA EST yk240f8.3 comes from this gene; cDNA EST yk387c8.3 comes from this gene; cDNA ES...		UNCLASSIFIED	264758
1625	95354748 (3249, 3250)	Novel Protein sim. GBank gl 4589822 db BAA76833.1 - (AB023206) KIAA0989 protein [Homo sapiens]		kinase	264489, 22278994, 22278995, 22278998, 35696286, 22278997, 22278998, 22278999, 264092, 264259, 29331824, 29331825, 29331827, 29331828, 264102, 264106, 264508, 33657084, 265017, 265018, 18108351, 264683, 264368, 264288, 21906765, 21906768, 21906767, 21906769, 35695917, 265021, 264691, 65274820, 18108368, 263972, 18108376, 35896423, 264631, 264634, 22279000, 22279002
1626	94734369 (3251, 3252)	Novel Protein sim. GBank gl 587907 gb AA046844.1 AF16090 - (AF160904) BcDNA_HL05936 [Drosophila melanogaster]			52844507, 52845156, 52846365, 52846842, 22278994, 56994075, 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 68714117, 29331826, 60432289, 29331827, 29331828, 35696052, 264906, 66712502, 264909, 265008, 265009, 60432228, 60433358, 60433438, 21908754, 52846317, 52644288, 265011, 87168559, 264604, 265018, 264448, 264369, 264288, 264768, 52844229, 264689, 21906765, 21906768, 35895917, 265021, 265022, 52644150, 33657023, 65274820, 27486261, 27486262, 27486285, 35695783, 263972, 52844332, 60170394, 87168516, 60432113, 264567
1627	83368773 (3253, 3254)			UNCLASSIFIED	264288
1628	85708459 (3255, 3256)	Novel Protein sim. GBank gl 3668087 (AC004667) - hypothetical protein [Arabidopsis thaliana]		eph	264288, 264686, 264767, 22279002

1629	84993841 (3257, 3258)	Novel Protein sim. GBank gi 4240175 dbj BAA74866.1 - (AB020850) KIAA0843 protein [Homo sapiens]		264555	
1630	87779027 (3259, 3260)			29331822, 29331827, 265010, 264693, 264634, 22279002	UNCLASSIFIED
1631	87758454 (3261, 3262)	Novel Protein sim. GBank gi 1915892 emb CAA69995 - (Y08740) tom-1A protein [Gallus gallus]		55811957, 264259, 33657023, 264693, 29331822, 29331824, 29331827, 29331828, 264906, 264908, 55811578, 264910, 264634, 264636, 264637, 56182323, 264559, 264758, 18108365, 264563, 264764, 264768	UNCLASSIFIED
1632	87871692 (3263, 3264)	Novel Protein sim. GBank gi 255850 dbj BAA22886 - (D83850) hepatoma-derived growth factor [Mus musculus]		264687, 264769, 264691, 264692, 2914849, 264509, 264905, 264907, 264511, 264512, 264482, 264681, 264763, 264682, 264683	UNCLASSIFIED
1633	87773683 (3265, 3266)			264486, 264259, 264907, 264808, 264909, 264628, 264629, 264631	
1634	85992817 (3267, 3268)	Novel Protein sim. GBank gi 4887229 gb AAD32244.1 AF150755 - (AF150755) microtubule-actin crosslinking factor [Mus musculus]	Contains protein domain (PF00435) - spectrin repeat	265007, 264637, 22279002	
1635	84232600 (3269, 3270)				
1636	80413227 (3271, 3272)	Novel Protein sim. GBank gi 455751 ref NP_001339.1 pDAPK - death-associated protein kinase 3		65274572, 22278998, 35696052, 52644045, 264511, 265008, 265009, 265010, 265011, 265018, 265019, 264448, 264369, 21906785, 21906788, 265021, 264690, 264482	UNCLASSIFIED
1637	80070435 (3273, 3274)	Novel Protein sim. GBank gi 3420051 AC004660 - unknown protein [Arabidopsis thaliana]	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	22278995, 264594, 264763, 265020, 264558	UNCLASSIFIED
1638	87101854 (3275, 3276)			21906765, 21906767, 22278998, 35696288, 22278998, 264259, 264692, 264693, 29331824, 33657109, 264508, 264906, 18108370, 264629, 265007, 33657402, 21906754, 264602, 264604, 264764, 264683, 264568, 264288	
1639	84322194 (3277, 3278)	Novel Protein sim. GBank gi 5420389 emb CAB46680.1 - (AJ243460) proteophosphoglycan [Leishmania major]		264488, 18108394, 65274572, 56182575, 35696286, 29331824, 29331826, 29331827, 35696052, 264908, 56182435, 264112, 265008, 265009, 264757, 264758, 55811386, 264603, 264760, 18108351, 264764, 264288, 264766, 264768, 21906767, 55811957, 264691, 33657023, 65274620, 18108370, 55810764, 55811576, 264558, 264639, 83373044, 18108385, 87168518	UNCLASSIFIED

1640	94143185 (3279, 3280)	Novel Protein sim. GBank gij2842469jemb CAA16847.1 - (AL021747) hypothetical protein [Schizosaccharomyces pombe]		UNCLASSIFIED	58182575, 58994075, 35696286, 60432049, 60432289, 29331827, 35698052, 52844045, 58182435, 284510, 285008, 285007, 285008, 284910, 285009, 33857402, 55812038, 285010, 285011, 285017, 285018, 284288, 52844229, 21908785, 21908768, 21908768, 35895917, 265021, 60170615, 52844150, 33657023, 33857108, 33857349, 18108374, 35898423, 65274781, 35895855, 284632, 284555, 58182323, 22279000
1641	87625160 (3281, 3282)			UNCLASSIFIED	29148489, 285008, 285007, 285008, 285009, 55812038, 285010, 285011, 284555, 284556, 284558, 18108383
1642	94312557 (3283, 3284)	Novel Protein sim. GBank gij1575333 (U60416) - myr 6 myosin heavy chain [Rattus norvegicus]	Contains protein domain (PF01843) - struct DIL domain		22278999, 28147820, 29331828, 29331828, 33658970, 55812038, 285010, 285018, 285019, 18108351, 284689, 285020, 285022, 284690, 33657023, 65274620, 35695783, 52844332, 18108381, 60170394, 58182323, 18108388, 87188518, 22279002, 284584, 29331825, 29331827, 29331828, 21908754, 285019, 284288, 284693, 33857349, 18108370, 18108376, 284555, 83373044, 22279002, 264482
1643	94131768 (3285, 3286)				284905, 284907, 284908, 284910, 285009, 284757, 284758, 284761, 284762, 284763, 284766, 284768, 284769, 284628, 284629, 284630, 284631, 284632, 284563, 284584, 284585, 284586, 284587
1644	88085125 (3287, 3288)			UNCLASSIFIED	284685, 284693
1645	95013858 (3289, 3290)	Novel Protein sim. GBank gij1076802 pir S49915 - extensin like protein - maize		UNCLASSIFIED	22278994, 58994075, 35696286, 284258, 29331824, 29331825, 29331826, 60432289, 284508, 60433358, 60433438, 87168559, 285018, 284687, 35895917, 284692, 33657023, 33657182, 27486281, 27486265, 33857349, 60432113, 284583, 284584
1646	95362691 (3291, 3292)			UNCLASSIFIED	29331822, 284908, 284908, 284369, 21908788, 60170815, 284639, 22278000
1647	94278428 (3293, 3294)	Novel Protein sim. GBank gij5002573jemb CAB44338.1 - (Y17460) alpha-N-acetylgalactosamine alpha-2,6-sialyltransferase [Fugu rubripes]		UNCLASSIFIED	285009, 284686, 55811957, 35895917, 55810764, 284558, 58182323, 284558, 18108385
1648	87642098 (3295, 3296)		Contains protein domain (PF00086) - Zinc finger, C2H2 type	UNCLASSIFIED	

1649	95347628 (3297, 3298)	Novel Protein sim. GBank gl 854065 emb CAA58337 - (X83413) U88 [Human herpesvirus 6]		cadherin	264488, 22278995, 35896286, 22278998, 22278997, 22278999, 60432049, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 264905, 264907, 66712502, 264908, 52644045, 264909, 56182435, 264511, 265007, 265008, 265009, 264591, 264593, 60433438, 264596, 55812038, 21908754, 265011, 264601, 264602, 265017, 265018, 265019, 264682, 264448, 264784, 264683, 264288, 264768, 264685, 264687, 264768, 264688, 264769, 52644229, 264689, 21908765, 21908768, 21908767, 21908768, 55811957, 35695917, 265021, 265022, 52644150, 264692, 33657023, 33657109, 20281149, 18108370, 264628, 18108374, 18108376, 35698423, 35695855, 264632, 264634, 264635, 264636, 18108380, 264639, 264558, 18108382, 18108384, 18108385, 18108387, 264080, 264404, 60432113, 22278000, 22278002, 264482, 264585, 264566, 264487, 265011, 264602, 21908767, 18108374, 18108377, 18108385
1650	87418539 (3289, 3300)	Novel Protein sim. GBank gl 3647335 emb CAA21059 - (AL031644) possible zinc-finger protein (Schizosaccharomyces pombe)			
1651	91839773 (3301, 3302)	Novel Protein sim. GBank gl 4884278 emb CAB43247.1 - (AL050037) hypothetical protein [Homo sapiens]		synthase	264488, 52645156, 18108397, 35698288, 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 264508, 264908, 29331830, 264910, 60432229, 21908754, 265010, 265011, 265017, 265019, 264448, 18108354, 264288, 264688, 21908765, 21908768, 21908769, 265022, 264692, 264693, 264629, 35695855, 264558, 264637, 264557, 264559, 83373044, 56526486, 22278000, 22278002, 264584
1652	86598622 (3303, 3304)	Novel Protein sim. GBank gl 1657837 (U73200) - p116Rip [Mus musculus]	Contains protein domain (PF00169) - PH domain	struc1	22278997, 29148498, 58182435, 21908754, 264368, 21908765, 21908768, 21908769, 265020, 52644150, 33657108, 22278000, 22278002
1653	94255993 (3305, 3306)	Novel Protein sim. GBank gl 3776054 emb CAA06273 - (AJ004999) Tapasin [Gallus gallus]	Contains protein domain (PF00047) - Immunoglobulin domain	glycoprotein	18108398, 22278995, 22278998, 264259, 29331822, 29331824, 60432289, 29331826, 29331827, 29331830, 264908, 265008, 265009, 60432229, 60433356, 60433438, 21908754, 265017, 265019, 264448, 264683, 264288, 265021, 265022, 264692, 18108384, 65274781, 18108384, 60432113, 264587, 33657109, 264585
1654	78758471 (3307, 3308)			UNCLASSIFIED	

1655	86689346 (3308, 3310)	Novel Protein sim. GBank gij3355717[embjCAA73496] - (Y13053) seryl-tRNA synthetase [Zea mays]		synthase	5284507, 35696286, 22278998, 22278999, 29331824, 29331825, 29331828, 33656970, 294908, 5264045, 264511, 264910, 52646317, 264288, 52644229, 33657023, 33657109, 52644332, 264557, 56182323, 56526486, 60432113
1656	79962287 (3311, 3312)	Novel Protein sim. GBank gij1890141[dbjBAA18947] - (D83206) P24 protein [Mus musculus]		UNCLASSIFIED	29331822, 29331824, 29331825, 264563
1657	87771894 (3313, 3314)	Novel Protein sim. GBank gij4557645[reijNP_001524.1]pHNR - heterogeneous nuclear ribonucleoprotein L	Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	dna_ma_bind	29331827, 265009, 21906766, 21906767, 265020, 265022, 33657109, 264638, 56526486, 264482
1658	87773778 (3315, 3316)	Novel Protein sim. GBank gij3877072[embjCAA87060] - (Z46937) similarity with ribosomal protein L21 [Caenorhabditis elegans]	Contains protein domain (PF00829) - Ribosomal prokaryotic L21 protein	UNCLASSIFIED	52846365, 35696286, 22278998, 22278999, 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 29146498, 264905, 264908, 52844045, 265006, 60433356, 264757, 60433338, 21906754, 265011, 18108351, 264448, 264369, 264288, 264766, 264788, 21906765, 21906767, 21906768, 21906769, 29146629, 265021, 265022, 18108362, 263969, 263971, 18108374, 35696423, 18108383, 22279000, 264482
1659	88230101 (3317, 3318)	Novel Protein sim. GBank gij539218[pirjS38038] - hypothetical protein YKL201c - yeast (Saccharomyces cerevisiae)		UNCLASSIFIED	52846317, 21906768, 21906767, 21906768, 87168518, 22278998, 265020, 22278999, 87168559, 264603, 265017, 264631, 265018, 265019, 22279002, 264482, 264635, 264585
1660	94315313 (3318, 3320)	Novel Protein sim. GBank gij2497012[spjQ10010]YSA4 CAEEL - HYPOTHETICAL 28.6 KD PROTEIN T19C3.4 IN CHROMOSOME III		UNCLASSIFIED	264488, 35696286, 264259, 35696052, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 265006, 265007, 264512, 265008, 264910, 264592, 264598, 265010, 264600, 264602, 265017, 265018, 264605, 264780, 264764, 264288, 264766, 264688, 264788, 264769, 264689, 21906766, 35695917, 264690, 33657023, 264693, 33657109, 264629, 35696423, 35695855, 264634, 264635, 264555, 264638, 264637, 264556, 264638, 264639, 264559, 18108385, 18108388, 264563, 264483, 264564, 264565, 264566, 264486, 264567

1661	94234071 (3321, 3322)	Novel Protein sim. GBank gi 4759100 ref NP_004759.1 pSFRS - splicing factor, arginine/serine-rich 11	Contains protein domain (PF00076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain)	UNCLASSIFIED	264488, 22278998, 264259, 29331824, 29331826, 29331827, 29331828, 264509, 68712502, 29331830, 264908, 52644045, 265007, 264512, 60433358, 60433438, 55812038, 21908754, 265019, 264448, 264766, 264768, 264769, 21908768, 21908769, 265020, 33857023, 33657109, 85274791, 87168518, 264482, 264583, 264584, 264585, 264587
1662	94135172 (3323, 3324)	Novel Protein sim. GBank gi 1730502 sp P52875 PF27_MOUSE - TRANSMEMBRANE PROTEIN PF27			18108392, 29331822, 29331828, 20281100, 264108, 265008, 265007, 265008, 18108348, 21908766, 18108385, 18108366, 18108374, 83373044, 18108385
1663	94217146 (3325, 3326)	Novel Protein sim. GBank gi 4884136 emb CAB43275.1 - (AL050107) hypothetical protein [Homo sapiens]	Contains protein domain (PF00397) - WW domain	kinase	52645156, 56182575, 22278994, 22278995, 35698286, 22278998, 56984075, 22278997, 22278998, 22278999, 264259, 29331822, 29331826, 29331827, 29331828, 33658970, 29331830, 264908, 56182435, 264511, 60433356, 33857402, 33108954, 87168474, 87168559, 265017, 265018, 264605, 18108351, 264764, 264288, 264766, 264768, 21908765, 21908768, 21908767, 21908768, 21908769, 265021, 265022, 264691, 33857023, 264693, 263967, 33657109, 264630, 52644332, 83373044, 87168518, 60432113, 22278000
1664	94234076 (3327, 3328)	Novel Protein sim. GBank gi 3043692 dbj BAA25510 - (AB011156) KIAA0584 protein [Homo sapiens]		UNCLASSIFIED	264488, 263994, 35698288, 29331824, 35698052, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 265009, 264910, 60170831, 264591, 264592, 264595, 87168474, 265011, 264600, 264601, 264604, 264605, 264760, 264762, 18108351, 264681, 264682, 264763, 264683, 264764, 264288, 264684, 264766, 264687, 264768, 264769, 21908764, 21908765, 21908767, 35695917, 265021, 264534, 60170815, 264690, 264691, 264692, 33657109, 33657182, 264628, 18108370, 264628, 35698423, 35695855, 264634, 264635, 264555, 264638, 264637, 264638, 264639, 264558, 83373044, 87168518, 264583, 264586, 264488

1665	91226852 (3329, 3330)	Novel Protein sim. GBank gl 1083506 p j S50085 - siadtheadin - mouse	Contains protein domain (PF00047) - Immunoglobulin domain	immunoglob	264468, 26331826, 29331828, 264509, 264906, 264907, 264909, 264510, 264511, 264910, 264592, 264593, 264595, 264758, 264596, 264800, 264760, 264762, 264764, 264766, 264768, 264829, 264630, 264634, 264636, 83373044, 264564, 264566, 264567, 264488
1666	95358160 (3331, 3332)	Novel Protein sim. GBank gl 3913431 sp O42843 DDX8_SCHPO - PUTATIVE PRE-MRNA SPLICING FACTOR ATP-DEPENDENT RNA HELICASE SPAC10F8.02C	Contains protein domain (PF00575) - S1 RNA binding domain	helicase	56994075, 22278999, 264259, 29331824, 29331826, 29331827, 29146498, 265009, 33109954, 87168559, 265018, 264288, 264688, 21908787, 21908789, 264691, 33857182, 18108370, 18108374, 18108385, 22279002
1667	91226855 (3333, 3334)	Novel Protein sim. GBank gl 5689535 db BAA83051.1 - (AB029022) KIAA1099 protein [Homo sapiens]	Contains protein domain (PF01412) - Putative GTP-ase activating protein for Arf	transport	264259, 29331822, 29331826, 264905, 264906, 264908, 264510, 265009, 264595, 264758, 265011, 87168559, 265017, 265018, 265019, 264448, 264766, 264888, 21908785, 21908767, 21908769, 265020, 265021, 60170815, 264690, 264692, 264693, 18108368, 18108370, 263972, 55810764, 264555, 83373044, 80432113, 22279000, 22279002
1668	88095135 (3335, 3336)	Novel Protein sim. GBank gl 2076894 gb AAB53983.1 - (AF002197) short region of weak similarity to protein kinase C; contains similarity to Pfam domain PF00130 (DAG_PE_bind). Score=10.0, E-value=0.0034, N=1 [Caenorhabditis elegans]		kinase	66714117, 264508, 264509, 264906, 264907, 264908, 264511, 264910, 264764, 264887, 264889, 33657109, 35696423, 35695855, 264632
1669	91227846 (3337, 3338)	Novel Protein sim. GBank gl 3875371 emb CAA85414.1 - (Z36948) contains a valine and arginine rich domain, possesses weak similarity with the RNA binding domains from RNA splicing factor U2AF 65 KO subunit; cDNA EST EMBL:D84658 comes from this gene; cDNA EST EMBL:D68828 comes fr...		UNCLASSIFIED	29331825, 33109954, 264369, 264787, 264689, 33657109, 83373044
1670	87628008 (3339, 3340)			UNCLASSIFIED	264259, 29331824, 29331827, 60433438, 265022, 264636
1671	87346372 (3341, 3342)	Novel Protein sim. GBank gl 482451 sp P34244 KKK1_YEAST - PROBABLE SERINE/THREONINE-PROTEIN KINASE YKL101W	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	kinase	35696288, 22278997, 29331825, 264909, 21908754, 265017, 265018, 265019, 264682, 264883, 264768, 264888, 21908766, 21908767, 21908768, 21908769, 264691, 264555, 264556, 22279000, 264566
1672	86291834 (3343, 3344)	Novel Protein sim. GBank gl 1814270 (U74586) - double-stranded RNA specific adenosine deaminase [Rattus norvegicus]		deaminase	264906, 264909, 264632, 18108381

1673	88095137 (3345, 3346)	Novel Protein sim. GBank gij2076894[jb]AAB3383.1] - (AF002197) short region of weak similarity to protein kinase C; contains similarity to Plam domain PF00130 (DAG_PE_bind), Score=10.0, E-value=0.0034, N=1 [Caenorhabditis elegans]	Contains protein domain (PF00130) - Phorbol esters/diacylglycerol binding domain (C1 domain)	Kinase	264488, 264569, 18108394, 56994075, 22278998, 284259, 35896052, 284508, 284905, 264509, 264908, 264907, 264908, 264909, 264510, 264910, 60170831, 264592, 284594, 264595, 264758, 264601, 264760, 264762, 264683, 264764, 264288, 264766, 264886, 264768, 264687, 264769, 264689, 264690, 33857023, 264692, 264693, 33857109, 264628, 264629, 18108374, 35898423, 35898555, 264631, 264632, 264634, 264635, 264637, 264558, 264638, 264639, 264563, 264482, 264564, 264565, 264566, 264567, 264488
1674	88258028 (3347, 3348)	Novel Protein sim. GBank gij526246[jmb]CAB45893.1] - (AL080082) hypothetical protein [Homo sapiens]		Kinase	29331822, 29331824, 284908, 52644045, 60433358, 87168559, 264448, 264288, 264888, 264691
1675	87608468 (3349, 3350)	Novel Protein sim. GBank gij3128366 (AF010496) - 50S ribosomal protein 19 [Rhodobacter capsulatus]		UNCLASSIFIED	56181886, 35896286, 22278997, 22278998, 284259, 29331824, 29331827, 35896052, 66712502, 264764, 264288, 264888, 264687, 35895917, 265020, 264690, 264693, 35895763, 18108370, 35898423, 35895855, 264637, 264639, 18108385, 264564
1676	95358086 (3351, 3352)	Novel Protein sim. GBank gij4184085[jb]AAD05327] - (AF111091) latrophilin 3 splice variant bbaf [Bos taurus]		UNCLASSIFIED	284259, 29331827, 29331828, 264106, 264907, 265009, 264600, 265018, 264288, 21908785, 265020, 265022, 35895855, 83373044, 18108385
1677	87408587 (3353, 3354)	Novel Protein sim. GBank gij3327046[jb]BAA31591] - (AB014516) KIAA0816 protein [Homo sapiens]		UNCLASSIFIED	284908
1678	88868829 (3355, 3356)	Novel Protein sim. GBank gij550452 (U08469) - 3-methylcrotonyl-CoA carboxylase, biotin-carrier domain [Glycine max]	Contains protein domain (PF00289) - Carbamoyl-phosphate synthase (CPSase)	UNCLASSIFIED	29331824, 264102
1679	81214108 (3357, 3358)			carboxylase	284488, 18108392, 18108394, 52646842, 18108397, 18108398, 35896286, 29331824, 265006, 265007, 265008, 265009, 18108348, 265011, 18108351, 264683, 18108354, 18108358, 18108359, 21908765, 29148627, 29148829, 264690, 18108381, 18108382, 18108384, 18108385, 18108386, 264628, 18108378, 35898423, 35895855, 264635, 18108381, 18108382, 18108383, 18108384, 18108385, 18108388
1680	91005372 (3359, 3360)	Novel Protein sim. GBank gij2394478 (AF024500) - No definition line found [Caenorhabditis elegans]		transport	85274572, 22278994, 22278999, 66714117, 29331827, 56182435, 21908754, 265018, 264288, 21908769
1681	94324150 (3361, 3362)	Novel Protein sim. GBank gij5689537[jb]BAA83052.1] - (AB029023) KIAA1100 protein [Homo sapiens]		UNCLASSIFIED	22278996, 29331822, 284908, 264593, 264604, 265019, 264683, 55811957, 264690, 33857023, 35898423, 83373044, 264563
1682	86042710 (3363, 3364)			UNCLASSIFIED	284908, 265017, 264605

1683	94316213 (3365, 3366)	Novel Protein sim. GBank gi15031171 ref NP_005704.1 pGPBP - goodpasture antigen- binding protein	Contains protein domain (PF01852) - START domain	UNCLASSIFIED	263994, 35696286, 35696032, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264511, 265008, 265007, 265009, 264910, 264593, 264758, 265010, 265011, 264760, 264761, 264762, 264763, 264764, 264369, 264766, 264768, 35695917, 264692, 33657109, 35698423, 35695855, 264634, 264635, 264636, 264638, 264639, 83373044, 264486
1684	80053409 (3367, 3368)	Novel Protein sim. GBank gi1255371 (U53147) - coded for by C. elegans cDNA yk34a9.5; coded for by C. elegans cDNA yk34a9.3; Similar to guanylate kinase. [Caenorhabditis elegans]	Contains protein domain (PF00825) - Guanylate kinase	UNCLASSIFIED	264563, 264566 60424178, 52846842, 22278984, 35696286, 22278998, 264259, 52645080, 29331824, 29331826, 265007, 33657084, 265018, 264681, 264448, 264683, 264369, 264689, 21908765, 21908787, 21908788, 21908789, 265021, 264692, 85274620, 33657109, 27486262, 264635, 52644332, 56182323, 22278000
1685	94323182 (3368, 3370)	Novel Protein sim. GBank gi1255371 (U53147) - coded for by C. elegans cDNA yk34a9.5; coded for by C. elegans cDNA yk34a9.3; Similar to guanylate kinase. [Caenorhabditis elegans]	Contains protein domain (PF00825) - Guanylate kinase	UNCLASSIFIED	264563, 264566 60424178, 52846842, 22278984, 35696286, 22278998, 264259, 52645080, 29331824, 29331826, 265007, 33657084, 265018, 264681, 264448, 264683, 264369, 264689, 21908765, 21908787, 21908788, 21908789, 265021, 264692, 85274620, 33657109, 27486262, 264635, 52644332, 56182323, 22278000
1686	87820710 (3371, 3372)	Novel Protein sim. GBank gi12244707 dbj BAA21115.1 - (AB005287) thrombospondin 1 [Bos taurus]	UNCLASSIFIED	UNCLASSIFIED	59182575, 264259, 264508, 264905, 264509, 264907, 264908, 264510, 264511, 265006, 264512, 265009, 264910, 264758, 265010, 265011, 264605, 18108351, 264764, 264766, 18108357, 264768, 18108362, 264628, 264630, 264631, 264634, 264635, 264637, 264638, 264639, 264565, 264486, 264567, 35695917, 264906, 264907, 264908, 264510, 265008, 265007, 264910, 264558, 18108381, 18108383, 265011
1687	94719400 (3373, 3374)	Novel Protein sim. GBank gi14680879 gb AAD27729.1 AF13295 - (AF132954) CGI-20 protein [Homo sapiens]	UNCLASSIFIED	UNCLASSIFIED	264905, 264910, 264760, 264629, 264555 264569, 35696286, 22278999, 264508, 264908, 264909, 56182435, 33657402, 60433438, 55812038, 265017, 265018, 264448, 264764, 264288, 264688, 29148629, 35695917, 265020, 265021, 263972, 18108374, 65274791, 83373044, 264089, 264369, 21906766, 264692, 264639, 87168518
1688	82158442 (3375, 3376)	Novel Protein sim. GBank gi14240193 dbj BAA74875.1 - (AB020659) KIAA0852 protein [Homo sapiens]	UNCLASSIFIED	UNCLASSIFIED	264905, 264910, 264760, 264629, 264555 264569, 35696286, 22278999, 264508, 264908, 264909, 56182435, 33657402, 60433438, 55812038, 265017, 265018, 264448, 264764, 264288, 264688, 29148629, 35695917, 265020, 265021, 263972, 18108374, 65274791, 83373044, 264089, 264369, 21906766, 264692, 264639, 87168518
1689	94325049 (3377, 3378)	Novel Protein sim. GBank gi14240193 dbj BAA74875.1 - (AB020659) KIAA0852 protein [Homo sapiens]	UNCLASSIFIED	UNCLASSIFIED	264905, 264910, 264760, 264629, 264555 264569, 35696286, 22278999, 264508, 264908, 264909, 56182435, 33657402, 60433438, 55812038, 265017, 265018, 264448, 264764, 264288, 264688, 29148629, 35695917, 265020, 265021, 263972, 18108374, 65274791, 83373044, 264089, 264369, 21906766, 264692, 264639, 87168518
1690	83255346 (3379, 3380)	Novel Protein sim. GBank gi13800736 (AF031572) - seven- pass transmembrane receptor precursor [Mus musculus]	Contains protein domain (PF00028) - Cadherin domain	cadherin	264905, 264910, 264760, 264629, 264555 264569, 35696286, 22278999, 264508, 264908, 264909, 56182435, 33657402, 60433438, 55812038, 265017, 265018, 264448, 264764, 264288, 264688, 29148629, 35695917, 265020, 265021, 263972, 18108374, 65274791, 83373044, 264089, 264369, 21906766, 264692, 264639, 87168518
1691	88095223 (3381, 3382)	Novel Protein sim. GBank gi1273208 (AF039713) - No definition line found [Caenorhabditis elegans]	Cadherin domain		264905, 264910, 264760, 264629, 264555 264569, 35696286, 22278999, 264508, 264908, 264909, 56182435, 33657402, 60433438, 55812038, 265017, 265018, 264448, 264764, 264288, 264688, 29148629, 35695917, 265020, 265021, 263972, 18108374, 65274791, 83373044, 264089, 264369, 21906766, 264692, 264639, 87168518
1692	85106709 (3383, 3384)	Novel Protein sim. GBank gi121271 sp P02207 GLB_LAMFL - GLOBIN	Contains protein domain (PF00042) - Globin	UNCLASSIFIED	264108 26331826, 264508, 264905, 264907, 264595, 265010, 265011, 21908788, 33657023, 264629, 263978, 264558
1693	87012775 (3385, 3386)	Novel Protein sim. GBank gi121271 sp P02207 GLB_LAMFL - GLOBIN	Contains protein domain (PF00042) - Globin	UNCLASSIFIED	264108 26331826, 264508, 264905, 264907, 264595, 265010, 265011, 21908788, 33657023, 264629, 263978, 264558

1684	94208168 (3387, 3388)	Novel Protein sim. GBank gi15453932[re]NP_008225.1[pPOLR - polymerase (RNA) II (DNA directed) polypeptide J (13.3kD)]	Contains protein domain (PF01193) - RNA polymerases L / 13 to 16 kDa subunit	rnapolymerase	35696286, 22278998, 22278998, 22278998, 284259, 29331822, 29331825, 29331826, 28331828, 35896052, 29146499, 284903, 284908, 52844045, 264511, 265008, 265007, 265009, 284592, 60433356, 21906754, 265010, 265011, 18108351, 284763, 284682, 284448, 284683, 284288, 284768, 264689, 21906768, 60170615, 264691, 284692, 284693, 18108370, 18108374, 263978, 35898423, 35695855, 264558, 18108381, 18108385, 87168518, 284482, 284486 284634
1685	94719325 (3389, 3390)	Novel Protein sim. GBank gi14880679[gb]AAD27729.1[AF13295 - (AF132954) CGI-20 protein [Homo sapiens]		UNCLASSIFIED	
1686	87824038 (3391, 3392)	Novel Protein sim. GBank gi14220517[emb]CAA22990] - (AL035356) hypothetical protein [Arabidopsis thaliana]			22278997, 284259, 285010, 18108351, 284764, 21906766, 18108370
1687	85740963 (3393, 3394)	Novel Protein sim. GBank gi1505652 (U10382) - GP38b glycoprotein [Homo sapiens]		glycoprotein	284682
1688	87445285 (3395, 3396)	Novel Protein sim. GBank gi15052031[gb]AAD38411.1[AF15573 - (AF155739) axotrophin [Mus musculus]			58994075, 22278998, 22278999, 264509, 33657402, 284758, 87168474, 87168559, 265017, 265018, 284448, 284687, 28146627, 21906769, 28146829, 265020, 265022, 33657023, 284558, 87168518, 22278002 35896286, 284635
1689	87424783 (3397, 3398)	Novel Protein sim. GBank gi1543344[pi]IS41847 - zinc finger 5 protein - mouse	Contains protein domain (PF00098) - Zinc finger, C2H2 type	UNCLASSIFIED nucl_recp	28331824, 52644045, 265008, 265009, 263989, 263971
1700	87659161 (3399, 3400)	Novel Protein sim. GBank gi13877439[emb]CAA98652] - (Z72510) similarity to yeast UTR3 protein [Swiss Prot accession number P21374]; cDNA EST EMBL:D72822 comes from this gene; cDNA EST EMBL:D75763 comes from this gene; cDNA EST yk274e3.3 comes from this gene; cDNA EST yk274e3.5 c...		UNCLASSIFIED MHC	284092, 284110, 263977 22278995, 22278997, 284092, 29146498, 29146499, 264107, 284508, 284907, 284110, 284112, 285009, 60170831, 21906754, 265011, 265017, 284762, 18108351, 284288, 21906785, 35695917, 265021, 60170815, 263987, 33657109, 18108370, 263972, 263974, 18108374, 263976, 35895855, 264555, 263981, 60170394, 18108385, 58528486, 87168518, 60432113
1701	86570488 (3401, 3402)	Novel Protein sim. GBank gi1451544 (U04267) - proline-rich cell wall protein [Gossypium barbadense]		UNCLASSIFIED	284909, 265017, 284628, 264629, 284638
1702	87795092 (3403, 3404)	Novel Protein sim. GBank gi11263289 (U47856) - fibroin-4 [Araneus diadematus]		UNCLASSIFIED	284369
1703	79568651 (3405, 3406)	Novel Protein sim. GBank gi14519821[db]BAA75670.1] - (AB017614) OASIS protein [Mus musculus]			284569, 35696286, 284907, 265010, 284687, 284768, 284692, 284693, 284696, 284568 22278998, 22278998, 284259, 284509, 265018, 284764, 284685, 284686, 21906768, 21906769, 265022, 284691, 284558, 22279000
1704	86622978 (3407, 3408)	Novel Protein sim. GBank gi13123034[sp]Q15011[Y025_HUMAN - HYPOTHETICAL PROTEIN KIAA0025]			
1705	87795175 (3409, 3410)	Novel Protein sim. GBank			
1706	87790987 (3411, 3412)	Novel Protein sim. GBank			

1707	88041230 (3413, 3414)	Novel Protein sim. GBank gll4321684[gb AA015797] - (AF055470) ZNF258 [Homo sapiens]		UNCLASSIFIED	18108398, 22278997, 264259, 29147620, 29331826, 29146498, 284905, 284908, 285008, 264593, 264595, 284758, 284598, 265018, 284780, 18108351, 284764, 284766, 284689, 284693, 18108370, 35698423, 55811576, 264558, 87168518, 60432113, 264587
1708	91220519 (3415, 3416)	Novel Protein sim. GBank gll5174591[re NP_005947.1pMTHF - 5,10-methylenetetrahydrofolate dehydrogenase, 5,10-methylenetetrahydrofolate cyclohydrolase, 10-formyltetrahydrofolate synthetase	Contains protein domain (PF01268) - Formate--tetrahydrofolate ligase	synthase	56182575, 22278996, 56994075, 264259, 29331822, 29331824, 29331826, 29331827, 29331828, 29146498, 29146499, 29331830, 285009, 60170831, 33657402, 33108954, 87168559, 265018, 18108351, 264448, 21908765, 21908787, 21908768, 29148627, 29148629, 29148784, 60170615, 52644150, 33657023, 33657109, 18108374, 55811576, 264559, 18108385, 22279000, 264563
1709	80222583 (3417, 3418)			UNCLASSIFIED	264107, 55811957, 263974, 263976, 263977, 263981
1710	20754572 (3419, 3420)			UNCLASSIFIED	264556
1711	91013729 (3421, 3422)	Novel Protein sim. GBank gll5031735[re NP_005760.1pHEC- - N-acetylglucosamine 6-O-sulfotransferase		sulfotransferase	65274572, 29331824, 29331826, 284768, 60431528, 35698423, 60432113, 264583
1712	95330184 (3423, 3424)	Novel Protein sim. GBank gll5454168[re NP_006453.1pXAP4 - HBV associated factor	Contains protein domain (PF00841) - Zn-finger in Ran binding protein and others.	kinase	56994075, 264093, 284259, 29331822, 284098, 29331824, 29331827, 284107, 284110, 264511, 264592, 265011, 265018, 284683, 284686, 284689, 265020, 33657023, 263987, 33657109, 283974, 35698423, 35695855, 284630, 284636, 284558, 284568
1713	84143453 (3425, 3426)	Novel Protein sim. GBank gll160409 (M69183) - mature-parasite-infected erythrocyte surface antigen [Plasmodium falciparum]	Contains protein domain (PF00643) - B-box zinc finger.	UNCLASSIFIED	22278995, 264508, 284758, 18108351, 18108370, 283974, 18108374, 284634, 56182323, 83373044, 60432113
1714	87420048 (3427, 3428)				22278997, 284757, 21908765, 265020, 265021, 284692, 56526486
1715	84260257 (3429, 3430)	Novel Protein sim. GBank gll5689537[ob BAA83052.1] - (AB029023) KIAA1100 protein [Homo sapiens]		UNCLASSIFIED	284509, 284905, 284908, 284907, 284908, 284909, 284910, 284591, 265011, 264768, 284768, 284769, 284691, 284692, 284632, 284634, 284635, 284636, 284637, 284558, 284639, 284564
1716	87400449 (3431, 3432)	Novel Protein sim. GBank gll458946[ob BAA78761.1] - (AB012808) mBOCT [Mus musculus]		transport	56182575, 29331824, 60432289, 284109, 284909, 265007, 284600, 285019, 284888, 265020, 284693, 55811576, 264558, 60432113, 22278002

1717	87563223 (3433, 3434)	Novel Protein sim. GBank gi 2765411 emb CAA74749 - (Y14391) GTP-binding protein [Homo sapiens]	UNCLASSIFIED	264589, 264259, 29331825, 29331826, 29331828, 35696052, 264509, 264905, 264907, 264908, 264909, 264512, 265009, 264910, 264592, 264595, 264758, 264759, 265017, 264681, 264764, 264766, 264686, 18108357, 35695917, 264690, 264692, 264693, 264628, 264629, 35696423, 264630, 264631, 264635, 264636, 18108380, 264638, 264639, 18108388, 18108391
1718	87032628 (3435, 3436)	Novel Protein sim. GBank gi 2833262 sp Q14999 Y076_HUMAN - HYPOTHETICAL PROTEIN KIAA0076 (HA0936)	UNCLASSIFIED	265011, 264681, 264682, 264684, 264688, 264689, 21908765, 265021, 264691, 33657023, 264693, 18108370, 35695955, 264632, 264634, 264636, 18108386, 22278002
1719	94315258 (3437, 3438)	Novel Protein sim. GBank gi 4505197 ref NP_003473.1 pMLL2 - myeloid/lymphoid or mixed-lineage leukemia 2	UNCLASSIFIED	18108396, 65274572, 35696286, 22278997, 60432049, 58182181, 66714117, 60432289, 29331828, 35696052, 29331828, 264906, 29331830, 58182435, 264592, 60431735, 60433438, 55812038, 264759, 285010, 264600, 264601, 285017, 264448, 264764, 264288, 264769, 21908766, 21908769, 55811957, 285020, 265021, 52844150, 33657023, 33657109, 33657182, 27486282, 33657349, 35695763, 18108370, 60431528, 18108374, 35696423, 55811576, 35695955, 264631, 58182323, 264559, 264564, 264486, 56182375, 22278999, 264259, 29331824, 60432289, 29331827, 35696052, 264508, 264905, 264906, 264907, 264908, 264909, 264511, 264910, 264758, 21908754, 265011, 264601, 264760, 264762, 264288, 264766, 264686, 18108357, 264689, 21908765, 55811957, 264693, 20281149, 264629, 18108374, 55811576, 65274791, 264630, 20281071, 264634, 264635, 264636, 264637, 264556, 264638, 264639, 56182323, 87168518
1720	84653063 (3439, 3440)	Novel Protein sim. GBank gi 2129478 ref S51939 - chitinase (EC 3.2.1.14) precursor - beet	UNCLASSIFIED	22278994, 22278998, 29331822, 265006, 265007, 265008, 55812038, 21908754, 60174639, 265011, 87168559, 18108351, 18108354, 21908765, 21908766, 21908768, 21908769, 265020, 33657109, 18108370, 18108374, 264556, 60170394, 83373044, 18108385, 264486
1721	91722288 (3441, 3442)	Novel Protein sim. GBank gi 488646 emb CAB43381.1 - (AL050280) hypothetical protein [Homo sapiens]	UNCLASSIFIED	56994075, 29331824, 29331828, 265009, 18108351, 21908768, 265020, 33657023, 18108374, 83373044
1722	84134548 (3443, 3444)	Novel Protein sim. GBank gi 5689375 dbj BA82968.1 - (AB030644) tudor repeat associator with PCTAIRE 2 [Rattus norvegicus]	kinase Tudor domain	Contains protein domain (PF00587) - kinase Tudor domain

1723	95358181 (3445, 3448)	Novel Protein sim. GBank gij4426962[gblAAD20633] - (AF126082) Arf-like 2 binding protein BART1 [Homo sapiens]		UNCLASSIFIED	264488, 264687, 264769, 21906767, 21906768, 56182575, 55811957, 22278997, 22278998, 265020, 264259, 264692, 33657023, 29331822, 29331824, 29331825, 60432289, 33657182, 33658970, 33657349, 29146499, 264508, 264907, 18108370, 264829, 264908, 264909, 18108374, 55811576, 264510, 265008, 264511, 265007, 264910, 264632, 264591, 60432229, 264592, 60433356, 264594, 60433438, 264595, 83373044, 55812038, 33109954, 33657084, 87188518, 87188474, 265010, 265011, 87188559, 264600, 60432113, 264604, 265019, 264583, 264448, 264682, 264586, 264764, 264288, 264567, 264468, 264389, 264768
1724	87713806 (3447, 3448)	Novel Protein sim. GBank gij2340162 (AF005083) - dsRBP-ZFa [Xenopus laevis]	Contains protein domain (PF00086) - Zinc finger, C2H2 type	UNCLASSIFIED	264905, 18108359, 264693, 264628, 264631, 264638, 264555, 264558, 264558, 264559
1725	85655191 (3449, 3450)	Novel Protein sim. GBank gij3152662 (AF064604) - KE03 protein [Homo sapiens]	Contains protein domain (PF00023) - Ank repeat	homeobox	35696286, 264259, 29331822, 35696052, 264508, 264509, 264905, 264908, 264907, 264908, 264909, 264910, 265009, 264591, 264601, 264760, 18108351, 264681, 264764, 264288, 264766, 264768, 21906789, 35695917, 264628, 35696423, 264630, 264631, 264632, 264635, 264636, 264638, 87188518, 264588
1726	85754255 (3451, 3452)	Novel Protein sim. GBank gij4689348[gblAAD27861.1]AF13256 - (AF132562) BcDNA.LD14270 [Drosophila melanogaster]		UNCLASSIFIED	29148498, 264683, 264689
1727	85286382 (3453, 3454)			UNCLASSIFIED	264905, 265011, 264689, 21908768

1728	85349515 (3455, 3456)	Novel Protein sim. GBank gij4406549[gbjAA020027] - (AF131738) Unknown [Homo sapiens]		UNCLASSIFIED	60424179, 18108397, 58182575, 22278995, 58994075, 35898286, 22278997, 22278998, 22278999, 264094, 60432049, 264259, 28331822, 28331824, 58182181, 28331825, 60432289, 28331826, 28331827, 35898052, 284905, 284908, 284907, 28331830, 66712502, 284908, 58182435, 284511, 285008, 265009, 60432229, 60433356, 33857402, 60433438, 284759, 21908754, 87168474, 265010, 265011, 87168559, 265017, 285018, 265019, 5581150, 284681, 284448, 284682, 284763, 284683, 284288, 284684, 284369, 284685, 284786, 284687, 284789, 21908764, 284689, 21908765, 21908768, 21908767, 21908768, 35895917, 285020, 265021, 285022, 284535, 284691, 284692, 33857023, 284693, 33857109, 18108370, 284628, 283972, 284629, 18108374, 18108376, 55810764, 65274791, 35895855, 284631, 284634, 284635, 60431850, 284636, 284638, 60170394, 284638, 83373044, 58528486, 87188518, 60432113, 22279000, 22279002, 284584, 284586
1729	81227848 (3457, 3458)	Novel Protein sim. GBank gij854065[embjCAA58337] - (X83413) U88 [Human herpesvirus 8]		UNCLASSIFIED	284906, 284907, 284908, 284511, 284555, 83373044, 284596, 284588
1730	85483474 (3459, 3460)			UNCLASSIFIED	28331822, 28331825, 28331828, 284907, 284908, 284909, 285011, 284784, 284829
1731	86266088 (3461, 3462)	Novel Protein sim. GBank gij831600[pirjS47084 - hypothetical protein - rabbit]		UNCLASSIFIED	52648842, 284907, 284909, 56182435, 55811388, 87168559, 285018, 285019, 284760, 52844229, 55811576
1732	81218878 (3463, 3464)	Novel Protein sim. GBank gij4240231[dbjBAA74894.1] - (AB020878) KIAA0871 protein [Homo sapiens]		struct	58182575, 28331822, 28331824, 28331827, 68712502, 284591, 33857402, 60433356, 285019, 21908768, 21908769, 35895917, 285020, 265021, 284638, 56182323
1733	87617178 (3465, 3466)	Novel Protein sim. GBank gij1575756 (U70674) - m-Numb [Mus musculus]	Contains protein domain (PF00840) - synthase Phosphotyrosine interaction domain (PTB/PID)		284907, 284910, 33857402, 285010, 284681, 284683, 284684, 284888, 284769, 284691, 284692, 284693, 284628, 284638, 284558
1734	87785261 (3467, 3468)				284693

1735	88318838 (3469, 3470)	Novel Protein sim. GBank gi 4838807 gb AAD30566.1 AF146793 PFT27 [Mus musculus]	Contains protein domain (PF01169) - Uncharacterized protein family UPF0016		264488, 18108394, 18108398, 22278998, 60432049, 264259, 29331822, 264908, 265008, 265007, 265008, 265009, 264591, 33657402, 265010, 265011, 87168559, 265017, 18108351, 264682, 18108354, 264769, 264689, 21806765, 21906768, 21906769, 29148629, 29148784, 265021, 265022, 52644150, 18108384, 18108385, 33857109, 18108370, 18108374, 18108380, 18108385, 87168518, 264563, 18108390
1736	95362884 (3471, 3472)	Novel Protein sim. GBank gi 4885647 ref NP_005472.1 pTRAP - thyroid hormone receptor-associated protein complex component	UNCLASSIFIED		264490, 264259, 68714117, 68712502, 56182435, 285008, 265008, 264910, 60433358, 87168559, 265017, 265019, 18108351, 265020, 265022, 33657023, 33657109, 264555, 60431650, 264637, 60170394, 264558, 264639, 264564
1737	88165549 (3473, 3474)	Novel Protein sim. GBank gi 2143607 pir jS68695 - B/K protein - rat	Contains protein domain (PF00168) - C2 domain	kinase	29331822, 28331830, 264591, 265011, 265018, 265019, 22279002
1738	85788811 (3475, 3476)	Novel Protein sim. GBank gi 2225941 emb CAA697141 - (Y08460) Mdes protein [Mus musculus]	UNCLASSIFIED	UNCLASSIFIED	264908, 264909, 265008, 264910, 264566
1739	87328576 (3477, 3478)		Contains protein domain (PF00441) - Acy-CoA dehydrogenase	UNCLASSIFIED	35696052, 264603, 264557
1740	83592939 (3479, 3480)	Novel Protein sim. GBank gi 48091 emb CAA443091 - (X62452) YCR801 [Saccharomyces cerevisiae]		traffic	264604, 21806764, 18108364, 264629, 35695855, 264638
1741	95010100 (3481, 3482)	Novel Protein sim. GBank gi 4883898 gb AAD31695.1 AF13042 - (AF130420) serine protease-like protein isoform [Homo sapiens]	Contains protein domain (PF01363) - FYVE zinc finger	UNCLASSIFIED	60432289, 28331827, 264509, 265009, 60432229, 264759, 265017, 264767, 264688, 264689, 21806768, 285020, 33657109
1742	85788814 (3483, 3484)	Novel Protein sim. GBank gi 4505193 ref NP_003667.1 pMLD1 - membrane fatty acid (lipid) desaturase	UNCLASSIFIED	UNCLASSIFIED	264908, 264910, 264758, 265011, 264631, 264638, 264568
1743	86866475 (3485, 3486)				265017, 265020, 264692
1744	91224003 (3487, 3488)	Novel Protein sim. GBank gi 728832 sp P39189 ALU2_HUMAN - IIII ALU SUBFAMILY S8 WARNING ENTRY IIII	UNCLASSIFIED	UNCLASSIFIED	22278998, 264508, 264907
1745	20290075 (3489, 3490)		UNCLASSIFIED	UNCLASSIFIED	264558
1746	94326110 (3491, 3492)	Novel Protein sim. GBank gi 731758 sp P38873 YHY6_YEAST - HYPOTHETICAL 175.8 KD PROTEIN IN GND1-IK1 INTERGENIC REGION	Contains protein domain (PF00400) - WD domain, G-beta repeat	UNCLASSIFIED	65274791, 264639, 264559
1747	94324333 (3493, 3494)	Novel Protein sim. GBank gi 1658503 (U75467) - Alu [Drosophila melanogaster]		transcript factor	52846842, 29331824, 29331825, 52844045, 56182435, 265007, 52846317, 87168474, 265018, 18108351, 264369, 264768, 264689, 35695917, 60170815, 264691, 33657023, 65274791, 35695855, 60170394, 22279000, 264482

1748	88003580 (3495, 3496)	Novel Protein sim. GBank gi 450451 ref NP_001530.1 pHSJ2 - heat shock protein, DNAJ-like 2	Contains protein domain (PF00684) - eph DnaJ central domain (4 repeats)	264489, 56182575, 29331824, 56182435, 284112, 265007, 265019, 284764, 21908768, 265020, 264691, 55811576, 264635, 264555, 284556, 264557, 264559 264106
1749	83363091 (3497, 3498)	Novel Protein sim. GBank gi 5850780 gb AAD45948.1 AF151966 RGS protein RGS-17 [Gallus gallus]	Contains protein domain (PF00815) - Regulator of G protein signaling domain	33657402, 264288, 52644150, 263974, 83373044
1750	94321664 (3499, 3500)	Novel Protein sim. GBank gi 4956884 gb AAC28444.2 - (AF065184) hyperpolarization-activated, cyclic nucleotide- gated channel 2 [Homo sapiens]		
1751	83373058 (3501, 3502)	Novel Protein sim. GBank gi 2760161 dbj BAA24184 - (AB010054) outer arm dynein light chain 2 [Anthonodaris crassispina]	Contains protein domain (PF00560) - Leucine Rich Repeat	265010, 264369
1752	86456530 (3503, 3504)	Novel Protein sim. GBank gi 3915482 sp P74348 YG29_SYN3 - HYPOTHETICAL 35.0 KD PROTEIN SLR1829	Contains protein domain (PF00849) - RNA pseudouridylylase synthase	264510, 264593, 264682, 21908765, 18108370
1753	94235159 (3505, 3506)	Novel Protein sim. GBank gi 2852636 (AF007155) - unknown [Homo sapiens]	Contains protein domain (PF01553) - Acyltransferase	56994075, 22278996, 264908, 60170831, 264682, 284764, 264389, 264288, 264685, 264687, 21908766, 264692, 264693, 65274620, 65274781, 35695855, 264637, 264584
1754	88095323 (3507, 3508)	Novel Protein sim. GBank gi 731421 sp P3988 YEH4_YEAST - HYPOTHETICAL 53.3 KD PROTEIN IN HXT8-CAN1 INTERGENIC REGION	transport	264488, 35696286, 264509, 264806, 264807, 284908, 264909, 264511, 264910, 264591, 33657402, 264594, 264757, 264758, 284600, 264604, 264762, 264763, 264685, 264768, 264691, 264628, 35698423, 264632, 264634, 264637, 264638, 263981, 264558, 264639, 264563, 264584, 264585, 264568, 264567, 264686
1755	79470282 (3509, 3510)	Novel Protein sim. GBank gi 1176422 (U43184) - rhophilin [Mus musculus]	UNCLASSIFIED	
1756	92862614 (3511, 3512)	Novel Protein sim. GBank gi 432860 gb AAD20708 - (AC006300) putative glucose-induced repressor protein [Arabidopsis thaliana]		52646842, 22278994, 22278995, 56994075, 22278996, 22278997, 264259, 29331822, 60432289, 29331827, 33856970, 265006, 265009, 60432229, 60433356, 60433438, 33109954, 21908754, 265017, 265018, 265019, 264448, 264369, 264288, 21908765, 21908766, 21908767, 21908768, 21908769, 265020, 265021, 265022, 264692, 27488262, 27488264, 18108376, 20281152, 264558, 18108388, 87168518, 60432113, 22279000, 22279002, 264482
1757	95357380 (3513, 3514)	Novel Protein sim. GBank gi 5441615 emb CAB46856.1 - (AJ388557) zinc finger protein [Canis familiaris]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	22278997, 264259, 60432289, 29331827, 264908, 52644045, 265008, 264593, 265019, 264766, 21908768, 65274620, 18108385, 60432113, 264588, 264487

1758	87612971 (3515, 3516)	Novel Protein sim. GBank gij3881040[emb]CAA18403] - (AL021497) predicted using GeneFinder [Caenorhabditis elegans]		UNCLASSIFIED	22278994, 22278999, 29331822, 29331824, 29331825, 33658970, 264508, 265006, 285007, 265009, 284591, 33657402, 33109554, 87168474, 264600, 265017, 285018, 21908769, 265020, 265021, 33657023, 33657109, 284629, 18108374, 35895855, 284632, 52844332, 22279002, 284583
1759	36984372 (3517, 3518)			UNCLASSIFIED	284759
1760	87329718 (3519, 3520)	Novel Protein sim. GBank gij3262748[emb]CAB45888.1] - (AJ133120) Proline rich synapse associated protein 2 [Rattus norvegicus]		UNCLASSIFIED	56182575, 60432048, 35696052, 264805, 284908, 264907, 264908, 284909, 265008, 265009, 284910, 60432229, 284592, 284595, 55812038, 284758, 284782, 18108351, 284784, 264768, 264768, 284789, 21906765, 5581957, 35695817, 284690, 284692, 284628, 284629, 55811576, 35696423, 284632, 284634, 284638, 264557, 264639, 60432113, 284585, 284486
1761	87409586 (3521, 3522)	Novel Protein sim. GBank gij127749[sp]P10589[MYSC_ACACA - MYOSIN IC HEAVY CHAIN	Contains protein domain (PF00618) - Guanine nucleotide exchange factor for Ras-like GTPases; N-terminal motif	UNCLASSIFIED	29331822, 264910, 284685, 264686
1762	95319887 (3523, 3524)	Novel Protein sim. GBank gij3169158 (AC004770) - BC209730_2 [Homo sapiens]	Contains protein domain (PF00173) - Desaturase Heme-binding domain in cytochrome b5 and oxidoreductases		22278999, 264259, 284905, 264907, 60170831, 265010, 265011, 265017, 264448, 21908765, 21908768, 21908767, 21908768, 265021, 284690, 33657109, 18108374, 264558, 60170394
1763	91224013 (3525, 3526)	Novel Protein sim. GBank gij4809026[gb]AAD30082.1] - (AF132856) suppressor of G2 allele of skp1 homolog [Homo sapiens]			56181888, 29331825, 35696052, 264905, 284906, 264909, 264763, 284682, 284769, 35695917, 265022, 33657023, 18108374, 35696423, 284634
1764	87757697 (3527, 3528)	Novel Protein sim. GBank gij1360669[pir]ICGHU1V - collagen alpha 1(V) chain precursor - human	Contains protein domain (PF01410) - Fibrillar collagen C-terminal domain	collagen	22278998, 264259, 35696052, 29331828, 56182435, 265008, 265017, 265018, 284448, 284288, 21908766, 21906767, 29148627, 35695917, 284691, 33657023, 60432113, 22279002
1765	91230091 (3529, 3530)	Novel Protein sim. GBank gij486806[pir]IS35503 - finger protein neutralized - fruit fly (Drosophila melanogaster)		UNCLASSIFIED	284488, 18108394, 22278995, 22278997, 22278999, 284259, 264104, 264508, 284905, 284907, 264511, 265007, 265008, 284910, 265009, 21906754, 265010, 265017, 264603, 265018, 265019, 18108351, 284682, 284448, 284369, 284288, 264768, 18108358, 21908768, 21906767, 29148627, 29148629, 35695917, 265020, 265021, 264692, 284628, 264629, 18108374, 283976, 284636, 83373044, 22279000, 22279002

1766	95081201 (3531, 3532)	Novel Protein sim. GBank gi 2498087 sp Q09332 UGG_DROME - UDP- GLUCOSE:GLYCOPROTEIN GLUCOSYLTRANSFERASE PRECURSOR (DUGT)		glycoprotein	52845156, 87168559, 80170815, 33657023, 284693, 33657109, 27486281, 264555, 83373044
1767	87755998 (3533, 3534)	Novel Protein sim. GBank gi 170443 emb CAA18263.1 - (AL022238) dJ1042K10.4 (novel protein) [Homo sapiens]		UNCLASSIFIED	265017, 265019, 264686, 264768, 265020, 264692
1768	80253216 (3535, 3536)			UNCLASSIFIED	28331824, 28331825, 264591, 58182323
1769	87388988 (3537, 3538)			UNCLASSIFIED	264563
1770	85413144 (3539, 3540)	Novel Protein sim. GBank gi 115204 sp P06736 C1R_HUMAN - COMPLEMENT C1R COMPONENT PRECURSOR	Contains protein domain (PF00089) - Trypsin	complement	284488, 264768, 264769, 58182375, 55911957, 264680, 264691, 35696052, 284905, 264509, 264906, 264907, 264828, 264808, 264809, 264910, 264634, 264635, 264636, 264558, 264757, 264758, 55812038, 65274444, 264760, 264563, 264762, 264784, 264884, 264768
1771	84233542 (3541, 3542)	Novel Protein sim. GBank gi 3914181 sp P58558 OGT1_RAT - UDP-N- ACETYLGLUCOSAMINE-PEPTIDE N- ACETYLGLUCOSAMINYLTRANSFERASE 110 KO SUBUNIT (O-GLCNAC TRANSFERASE P110 SUBUNIT)	Contains protein domain (PF00515) - TPR Domain	transferase	264758, 264600, 264369, 55811957, 265020, 83373044, 22279000
1772	87643510 (3543, 3544)	Novel Protein sim. GBank gi 115204 sp P06736 C1R_HUMAN - COMPLEMENT C1R COMPONENT PRECURSOR	Contains protein domain (PF01528) - DHHC zinc finger domain		22278998, 28331828, 33109954, 265018, 265019, 264784, 21808785, 265020, 265021, 264558
1773	84116824 (3545, 3546)	Novel Protein sim. GBank gi 3978484 (AF085693) - G protein-coupled receptor kinase-associated ADP ribosylation factor GTPase-activating protein [Rattus norvegicus]	Contains protein domain (PF01412) - Putative GTP-ase activating protein for Arf	potassium_channel	65274572, 58182575, 22278998, 35696286, 22278999, 264259, 28331824, 60424289, 28331825, 60432289, 35696052, 264106, 264509, 264908, 264907, 28331830, 264908, 52644045, 264511, 265008, 265007, 265008, 60170831, 60433438, 264758, 55811388, 87188559, 265017, 264604, 265019, 55811150, 264288, 58181582, 264889, 21906768, 21908767, 21908768, 21908769, 55811857, 265020, 265022, 52844150, 264691, 33657023, 264692, 264693, 60431528, 35696423, 35695855, 264636, 58182323, 18108367, 58526486, 22279000, 22279002, 264583, 264584, 264585, 264586, 264587
1774	84232573 (3547, 3548)	Novel Protein sim. GBank gi 2495889 sp Q15034 Y032_HUMAN - HYPOTHETICAL PROTEIN KIAA0032	Contains protein domain (PF00415) - Regulator of chromosome condensation (RCC1)	UNCLASSIFIED	65274572, 58182575, 35696052, 55812038, 33109954, 21908754, 265017, 21908767, 21908768, 21908769, 265020, 264691, 264636, 58182323, 22279002

1775	85359330 (3549, 3550)	Novel Protein sim. GBank gij1468199[dbj][BAA09487] - (D50928) The KIAA0138 gene product is novel. [Homo sapiens]	UNCLASSIFIED	65274572, 56182575, 56994075, 35696286, 22278997, 29331822, 29331826, 60432289, 29331828, 35696052, 29331830, 66712502, 284828, 56182435, 284511, 265007, 265009, 60170831, 60432229, 60433438, 55812038, 21908754, 85658542, 87168559, 264601, 265017, 265018, 265019, 284762, 284448, 284288, 284889, 21908765, 21906768, 21908767, 21908768, 265020, 265021, 265022, 264691, 33657023, 264692, 33657109, 27486281, 33657349, 18108370, 18108377, 35698423, 55811578, 35698555, 284632, 284634, 284636, 284639, 58182323, 83373044, 56526486, 87168518, 60432113, 22279000, 22279002, 264482, 264486, 264910
1776	94133758 (3551, 3552)	Novel Protein sim. GBank gij4589676[dbj][BAA76857.1] - (AB023230) KIAA1013 protein [Homo sapiens]		
1777	87447171 (3553, 3554)	Novel Protein sim. GBank gij321993[sp][P87115]YOK9_SCHPO - HYPOTHETICAL 118.5 KD PROTEIN C20G8.09C IN CHROMOSOME I	nuc_recl	56994075, 29331826, 265008, 87168474, 265017, 265018, 284761, 55811550, 284784, 58181562, 284689, 21908765, 21906768, 21908769, 35695917, 264690, 33657023, 35695763, 60431528, 35696423, 55811578, 35695855, 22279000, 22279002, 264584, 29331826, 29331827, 35696052, 284512, 265007, 265009, 265017, 265019, 284782, 18108351, 284769, 21908765, 21908768, 21908767, 21908768, 21908769, 265020, 265021, 264691, 284693, 18108370, 284556, 83373044
1778	94851624 (3555, 3556)	Novel Protein sim. GBank gij387564[emb][CAA91454.1] - (Z66561) Similarity to Human rab13 protein (PIR Acc. No. A49847). Contains the ATP/GTP-binding site motif (PROSITE PS00017); cDNA EST EMBL:M89412 comes from this gene; cDNA EST yk212g9.3 comes from this gene; cDNA EST yk212g9...	UNCLASSIFIED	
1779	94133758 (3557, 3558)	Novel Protein sim. GBank gij4589676[dbj][BAA76857.1] - (AB023230) KIAA1013 protein [Homo sapiens]	UNCLASSIFIED	29148627, 35696286, 29147620, 265006, 265007, 265008, 18108385, 65274727, 264482, 264369, 284768
1780	87023497 (3559, 3560)		UNCLASSIFIED	Contains protein domain (PF00807) - Apidaech
1781	84047477 (3561, 3562)		UNCLASSIFIED	284508, 284908, 264639
1782	88094607 (3563, 3564)	Novel Protein sim. GBank gij729225[sp][P41237][CTXN_RAT] - CORTEXIN		284259, 29331822, 264508, 264905, 284908, 284907, 264908, 265007, 265009, 284910, 264591, 264758, 284764, 284288, 284768, 264769, 284635, 284638, 284637, 284639, 264593
1783	85717905 (3565, 3566)	Novel Protein sim. GBank gij2257543[dbj][BAA21436] - (AB004538) protein arginine N-methyltransferase [Schizosaccharomyces pombe]	interferon	264768

1784	95197093 (3567, 3568)	Novel Protein sim. GBank gll1755049 (U55042) - myosin X [Bos taurus]	Contains protein domain (PF00169) - struct PH domain	35898288, 284259, 35898052, 284508, 284905, 284908, 284907, 66712502, 284908, 284909, 285007, 285008, 285009, 284910, 284591, 284594, 284757, 284758, 284759, 205010, 285011, 284601, 284602, 284604, 284605, 18108351, 284762, 284763, 284764, 284369, 284768, 284687, 284768, 284688, 21908768, 35695917, 284690, 284691, 284692, 284693, 284628, 18108374, 35898423, 284631, 284632, 284635, 284637, 284638, 284639, 18108385, 22279000, 22279002, 284585, 284566, 284488
1785	95357475 (3569, 3570)	Novel Protein sim. GBank gll4589552[db][BAA76798.1] - (AB023171) KIAA0954 protein [Homo sapiens]	UNCLASSIFIED	65274572, 284259, 28331822, 28331824, 28331825, 28331827, 28331828, 284905, 284906, 284908, 66712502, 56182435, 284511, 285007, 60433356, 55811150, 284683, 284369, 284687, 52644229, 21908767, 52644150, 33857023, 65274620, 33657182, 65274791, 35895855, 284555, 65274727, 22279002
1786	85288485 (3571, 3572)	Novel Protein sim. GBank gll17788[sp]P28770[CYA4_RAT ADENYLATE CYCLASE, TYPE IV (ATP PYROPHOSPHATE-LYASE) (ADENYLYL CYCLASE)]	UNCLASSIFIED	284908, 35696423, 284636
1787	87434784 (3573, 3574)	Novel Protein sim. GBank gll3877175[emb][CAA90338.1] - (Z50028) cDNA EST yk321h8.5 comes from this gene; cDNA EST EMBL:D68896 comes from this gene; cDNA EST yk395f9.5 comes from this gene [Caenorhabditis elegans]	UNCLASSIFIED	284488, 284905, 284908, 284909, 284595, 284764, 284766, 284692, 60431528, 284629, 284636, 284584, 284586
1788	91228779 (3575, 3576)	Novel Protein sim. GBank gll2088669 (AF003130) - F55A12.9 gene product [Caenorhabditis elegans]	UNCLASSIFIED	284488, 83373044
1789	88094529 (3577, 3578)		UNCLASSIFIED	284488, 28331828, 284909, 18108351, 284288, 285021, 284555, 284638
1790	82489734 (3579, 3580)			35898052, 284905, 284906, 284907, 284908, 284909, 285008, 284910, 284758, 265011, 285019, 284764, 284766, 284769, 284628, 284635

1781	95197259 (3581, 3582)	Novel Protein sim. GBank gji2114321[dbj BAA20037] - (D88733) membrane glycoprotein [Equine herpesvirus 1]	Contains protein domain (PF00047) - Immunoglobulin domain	glycoprotein	284488, 284686, 284687, 284768, 18108394, 284769, 18108397, 284259, 284691, 284692, 33657023, 284693, 284509, 284905, 284908, 284628, 284907, 284629, 284908, 284909, 284510, 285008, 284511, 285008, 284630, 285009, 284631, 284910, 284632, 284634, 284635, 284555, 284636, 284592, 284637, 284593, 284638, 18108381, 284639, 284758, 285010, 285011, 284602, 22279000, 284604, 284760, 284564, 284681, 284782, 284565, 284763, 284683, 284566, 284764, 284288, 284684, 284567, 18108354, 18108391, 284685, 284766
1782	87792690 (3583, 3584)	Novel Protein sim. GBank gji4337106[gb AAD18082] - (AF129758) BAT4 [Homo sapiens]	Contains protein domain (PF01585) - G-patch domain	UNCLASSIFIED	22278997, 284259, 284508, 285007, 33657402, 87168559, 284369, 33657023, 35685855, 20281071, 284559, 18108387, 87168518
1783	95337877 (3585, 3586)	Novel Protein sim. GBank gji5579331[gb AAD45504.1][AF145732] endoplasmic reticulum alpha-mannosidase I [Homo sapiens]	Contains protein domain (PF01532) - Glycosyl hydrolase family 47	ATPase-associated	65274572, 22278995, 22278998, 22278997, 22278999, 284093, 284259, 29331824, 88714117, 60432289, 29331827, 29331828, 284103, 284105, 29331830, 265007, 264910, 285009, 60170831, 60433356, 21906754, 285010, 265017, 265019, 284681, 284682, 284288, 52644228, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 285021, 265022, 60170615, 52644150, 33657023, 33657109, 18108370, 18108374, 65274791, 20281071, 60432113, 22279000, 284482, 284564
1784	87759808 (3587, 3588)	Novel Protein sim. GBank gji4914604[emb CAB43877.1] - (AL050388) hypothetical protein [Homo sapiens]	Contains protein domain (PF01798) - Putative snRNA binding domain	UNCLASSIFIED	18108394, 22278995, 22278999, 264259, 29331822, 29331824, 29331825, 29148498, 29146499, 284508, 284905, 52644045, 284112, 265006, 265008, 264910, 60433356, 284757, 55812038, 87168474, 265011, 285017, 18108351, 284763, 264448, 284683, 284369, 21906765, 21906766, 21906767, 21906768, 29148784, 35695917, 60170815, 33657023, 284629, 18108374, 18108378, 35896423, 35895855, 284556, 284557, 284638, 284558, 18108385, 284584
1785	79747858 (3589, 3590)			UNCLASSIFIED	284632, 284635, 284638, 284595, 284596, 284607, 284568, 284909
1786	86598486 (3591, 3592)	Novel Protein sim. GBank gji595084[sp Q07803]EFGM_RAT - ELONGATION FACTOR G, MITOCHONDRIAL PRECURSOR (MEF-G)		glycoprotein	284488, 284907, 284909, 284594, 284595, 284768, 284687, 21906765, 21906767, 284628, 284630, 284559

1797	91223219 (3593, 3594)	Novel Protein sim. GBank gi1842111 (U87586) - decoy [Arabidopsis thaliana]		ribosomalprol	22278996, 22278997, 22278998, 22278999, 29331822, 264910, 60170831, 21908754, 52644229, 21908765, 21908768, 21908769, 35695917, 265022, 52644150, 264691, 33657023, 263967, 33657109, 22279000
1798	91221276 (3595, 3596)	Novel Protein sim. GBank gi2832505[db][BAA24608.1] - (D89340) dipeptidyl peptidase III [Rattus norvegicus]		peptidase	22278994, 56984075, 22278997, 22278998, 22278999, 264259, 29331826, 60432229, 29331828, 33656970, 265008, 60432229, 264757, 60433438, 21908754, 33657084, 87168559, 265017, 18108351, 264682, 264448, 264288, 21908765, 21908768, 21908767, 21908768, 21908769, 35695917, 265020, 265021, 33657023, 33657182, 27486261, 27486265, 33657349, 263973, 18108374, 55811576, 35695855, 18108385, 87168518, 22279000, 264488
1799	86321713 (3597, 3598)	Novel Protein sim. GBank gi5688541[db][BAA83054.1] - (AB029025) KIAA1102 protein [Homo sapiens]		eph	264908, 21908754, 21908767, 21908769, 265020, 33657023, 264692, 264693, 264404, 22279000
1800	87080118 (3599, 3600)	Novel Protein sim. GBank gi4680679[gb][AAD27729.1]AF132954 CGI-20 protein [Homo sapiens]		UNCLASSIFIED	284691, 264558, 264568
1801	95060723 (3601, 3602)	Novel Protein sim. GBank gi134920[sp]P21997[SSGP_VOLCA - SULFATED SURFACE GLYCOPROTEIN 185 (SSG 185)]			52644045, 265007, 264632
1802	87771012 (3603, 3604)	Novel Protein sim. GBank gi134920[sp]P21997[SSGP_VOLCA - SULFATED SURFACE GLYCOPROTEIN 185 (SSG 185)]			35686286, 66714117, 264508, 264509, 58182435, 264512, 18108351, 264888, 55611957, 264692, 55811576, 35695855, 264488
1803	95060725 (3605, 3606)	Novel Protein sim. GBank gi4880878[gb][AAD27729.1]AF132954 CGI-20 protein [Homo sapiens]			264686, 264488, 264687, 264489, 264768, 264769, 264689, 21908769, 35696286, 35695917, 264259, 264691, 264692, 264693, 20281099, 18108364, 35696052, 264508, 264509, 264805, 264908, 18108370, 264628, 264907, 66712502, 264908, 264909, 18108374, 18108376, 35696423, 35695855, 264510, 264511, 265008, 265007, 264512, 265008, 264910, 264631, 264632, 264634, 264635, 264591, 264636, 264637, 264592, 264638, 264593, 264639, 264594, 83373044, 264758, 264598, 18108385, 18108387, 265011, 264760, 264563, 18108351, 264762, 264564, 264448, 264565, 264763, 264683, 264764, 264586, 264288, 264486, 264567, 264765, 264768

1804	87770203 (3607, 3608)	Novel Protein sim. GBank gi 3879914 emb CAA98538.1 -(Z74043) predicted using GeneFinder; cDNA EST EMBL:C13850 comes from this gene; cDNA EST EMBL:C11575 comes from this gene; cDNA EST yk34314.5 comes from this gene [Caenorhabditis elegans]			52646365, 22278997, 22278999, 284905, 284908, 284909, 284910, 21906754, 284766, 21906765, 21906768, 35695917, 265020, 265022, 264691, 284637, 284639, 22279000, 264564, 264566
1805	95330375 (3609, 3610)	Novel Protein sim. GBank gi 545364 ref NP_006461.1 pEBBP - estrogen-responsive B box protein			29331824, 29331825, 29331826, 29331827, 29331828, 87188559, 284288, 284687, 52844229, 35698423, 264636, 50432113
1806	94133762 (3611, 3612)	Novel Protein sim. GBank gi 459676 db BAA76857.1 -(AB023230) KIAA1013 protein [Homo sapiens]	stncl		264094, 264105, 284908, 35698423, 285006, 265007, 285008, 264555, 264592, 265011, 265018, 284369
1807	86943032 (3613, 3614)				29331824, 284908, 284910, 33657023, 263978
1808	87642711 (3615, 3616)	Novel Protein sim. GBank gi 4884079 emb CAB43235.1 -(AL050008) hypothetical protein [Homo sapiens]	UNCLASSIFIED		284488, 35696286, 66714117, 35696052, 66712502, 284592, 80433438, 52844286, 265010, 284683, 264369, 284689, 55811957, 35695917, 33857109, 35695763, 55810764, 18108378, 35696423, 35695855, 56182323, 284563, 284584, 284487
1809	95321468 (3617, 3618)	Novel Protein sim. GBank gi 1916927 (U87965) - putative G protein [Mus musculus]	UNCLASSIFIED		284594, 55811150, 284686, 29148629, 29148784, 284690, 284629, 18108374, 284556, 264537, 264558
1810	88096316 (3618, 3620)	Novel Protein sim. GBank gi 1352944 sp P47179 Y9P_YEAST - HYPOTHETICAL 118.4 KD PROTEIN IN BAT2-DAL5 INTERGENIC REGION PRECURSOR	UNCLASSIFIED		284488, 35696052, 284905, 284906, 284907, 284908, 284909, 284511, 265009, 284910, 284592, 264593, 284594, 33657402, 284757, 284595, 284758, 284596, 284759, 284600, 284601, 264762, 284683, 264764, 264288, 284684, 284766, 284767, 284686, 284768, 284687, 284769, 284689, 265021, 284690, 284691, 284693, 284628, 284629, 18108374, 284630, 264631, 284632, 264634, 284635, 284636, 284637, 284638, 284639, 284583, 284586, 284486, 264567
1811	88086272 (3621, 3622)	Novel Protein sim. GBank gi 2134984 pir j37275 - death-associated protein kinase (EC 2.7.1.-) - human	Contains protein domain (PF00023) - kinase Ank repeat		284488, 264259, 284508, 284509, 284905, 284906, 284907, 56182435, 284511, 284512, 284910, 264758, 265011, 284600, 284604, 18108354, 284768, 284688, 284769, 264534, 60170815, 33857023, 284629, 284631, 284639, 284583, 284482, 284483
1812	78245772 (3623, 3624)				29331822, 28331824, 265019, 18108351, 21906769

1813	88090872 (3625, 3626)	Novel Protein sim. GBank gi5051638 gb AD36326.1 AF07372 - (AF07372) EH domain-binding mitotic phosphoprotein (Homo sapiens)	Contains protein domain (PF01417) - ENTH domain	glucoamylase	56182575, 264259, 29331824, 66714117, 29331828, 35696052, 284509, 264905, 284908, 264907, 284908, 66712502, 264909, 265007, 264910, 264591, 264593, 55812038, 285011, 265018, 264760, 264682, 264764, 284683, 264369, 264768, 264768, 264769, 21908768, 21908768, 284691, 264693, 18108374, 35695855, 264634, 264635, 264637, 264639, 284559, 22278000, 22278002, 264568
1814	88178047 (3627, 3628)	Novel Protein sim. GBank gi3843608 (AC005395) - hypothetical protein [Arabidopsis thaliana]		UNCLASSIFIED	284488, 35696286, 22278998, 264092, 264094, 264259, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 29331828, 264105, 264107, 52844045, 56182435, 265009, 60432229, 60433358, 87168474, 87168559, 264369, 264288, 21908765, 35695917, 285021, 285022, 33857023, 33657109, 18108374, 35698423, 264638, 56526486, 264482 22278999, 284508, 284509, 264907, 264908, 264910, 265011, 264760, 264766, 264634, 264636
1815	85296473 (3629, 3630)	Novel Protein sim. GBank gi117788 sp P26770 CYA4_RAT ADENYLATE CYCLASE, TYPE IV (ATP PYROPHOSPHATE-LYASE) (ADENYLYL CYCLASE)		struct	
1816	83738845 (3631, 3632)	Novel Protein sim. GBank gi1178623 sp P41946 YO88, CAEL - HYPOTHETICAL 93.9 KD PROTEIN T20812.6 IN CHROMOSOME III		UNCLASSIFIED	18108394, 18108397, 284509, 264907, 264908, 264909, 265009, 284591, 265011, 265017, 264687, 264689, 265022, 264691, 18108382, 18108368, 18108370, 18108374, 18108379, 264635, 284537, 264564, 264567 264488, 22278997, 22278998, 60432048, 29331822, 29331824, 60432289, 52844045, 60170831, 285017, 265018, 265019, 18108351, 264682, 52844229, 21906765, 21906767, 21906768, 52844150, 33657023, 33657109, 27486262, 18108370, 18108374, 60170394, 56182323, 22278002
1817	88095268 (3633, 3634)	Novel Protein sim. GBank gi3766377 emb CAA21429 - (AL031807) hypothetical protein [Schizosaccharomyces pombe]	Contains protein domain (PF00400) - IWD domain, G-beta repeat	histone	
1818	85806775 (3635, 3636)	Novel Protein sim. GBank gi3879121 emb CAA94370 - (Z70310) predicted using GeneFinder; Similarity to Mouse ankyrin (PIR Acc. No. S37771); cDNA EST EMBL: T01923 comes from this gene; cDNA EST EMBL: D32335 comes from this gene; cDNA EST EMBL: D32723 comes from this gene; cDNA EST E...	Contains protein domain (PF00023) - Ank repeat	transcript factor	35696286, 60433358, 264756, 264369, 264686, 21908769, 264693, 264632
1819	87759572 (3637, 3638)	Novel Protein sim. GBank gi5031865 ref NP_005771.1 pLHFP - Ipoma HMGIC fusion partner		UNCLASSIFIED	22278995, 22278997, 22278999, 264259, 29331826, 264908, 265007, 265008, 265009, 60432229, 265017, 265018, 265019, 264448, 264288, 21908768, 21908769, 265020, 18108361, 18108384, 22278000, 22278002, 264567

1820	87769455 (3839, 3840)				264905, 264907, 264594
1821	80431510 (3841, 3842)				264907, 264788, 263978
1822	91221523 (3643, 3644)	Novel Protein sim. GBank gi 4884130 emb CAB43272.1 - (AL050101) hypothetical protein (Homo sapiens)			22278995, 56994075, 22278996, 22278997, 22278998, 264259, 29331824, 29331825, 29331826, 35698052, 29331828, 264908, 29331830, 60170831, 264591, 264593, 60433356, 264598, 265017, 265019, 18108351, 264763, 264883, 21908765, 21908767, 21908768, 21908769, 35695917, 265020, 265021, 33657023, 18108364, 18108370, 35695855, 22278000, 22278002
1823	85522330 (3645, 3646)			UNCLASSIFIED	264488, 264259, 264511, 264288, 264768, 264693, 35698423, 264634, 18108385, 264488
1824	86612025 (3847, 3848)	Novel Protein sim. GBank gi 477072 pir A48018 - mucin 7 precursor, salivary - human	Contains protein domain (PF00047) - Immunoglobulin domain	UNCLASSIFIED	264907, 264908, 264909, 264511, 264631, 264634, 264635, 264637, 264638, 264639, 264758, 264588
1825	87430125 (3849, 3850)	Novel Protein sim. GBank gi 3036803 emb CAA18493 - (AL022373) hypothetical protein [Arabidopsis thaliana]		UNCLASSIFIED	60432049, 264910, 264487
1826	91723612 (3651, 3652)	Novel Protein sim. GBank gi 4680885 gb AAD27732.1 AF13295 - (AF132957) CGI-23 protein (Homo sapiens)		ATPase-associated	52844507, 52845156, 52646842, 22278994, 22278996, 56994075, 264259, 60432049, 52845080, 35698052, 68712502, 52844045, 265008, 265009, 60432229, 60433356, 80433438, 52846317, 52844296, 265011, 87168559, 264448, 264288, 264369, 264688, 52844229, 264689, 21908765, 21908768, 265020, 60170815, 52644150, 33657023, 27486262, 27486264, 27486265, 35695783, 35698423, 35695855, 83373044, 87168518, 264404, 22279002
1827	81647212 (3853, 3854)				264758

1828	95074017 (3655, 3656)	Novel Protein sim. GBank gi 4503571 ref NP_001419.1 PENO1 - endolase 1, (alpha)	Contains protein domain (PF00113) - Endolase	oncogene	264488, 52646842, 56182575, 22278896, 35696286, 22278897, 22278899, 264091, 264093, 60432049, 264259, 29331822, 29331824, 66714117, 29331825, 60432289, 29331826, 29331827, 29331828, 264105, 264508, 264907, 66712502, 52644045, 56182435, 265008, 264511, 264512, 265007, 265008, 265009, 60170831, 60432229, 264593, 60433358, 60433438, 264758, 33109954, 21908754, 87168474, 265010, 265011, 87168559, 265017, 265018, 264761, 264762, 264448, 264764, 264883, 264288, 264389, 18108355, 264768, 18108357, 18108358, 264688, 264769, 264689, 21908768, 21908769, 35695917, 265021, 60170815, 33657023, 33857349, 263972, 55811576, 35695855, 264635, 264555, 264558, 264638, 264557, 87168518, 22279000, 22279002, 264563, 264482, 264585, 264484, 264567
1829	80187720 (3657, 3658)				264508, 264634, 264509, 264482, 29331827, 264908, 265009, 264910
1830	94312942 (3659, 3660)	Novel Protein sim. GBank gi 2246532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma- associated herpesvirus]		nuclease	52645156, 22278894, 22278895, 35696286, 22278898, 22278897, 22278899, 22278899, 29331822, 29331825, 35688052, 52646317, 52644286, 87168559, 265018, 21908765, 21908766, 21908767, 21908768, 21908769, 35695917, 265021, 33657023, 52645129, 33837109, 33857182, 27488261, 27488262, 35695763, 263974, 35688423, 35695855, 52644332
1831	84138063 (3661, 3662)			UNCLASSIFIED	29331824, 35688052, 29331830, 264595, 264758, 265010, 265019, 265022, 264693, 65274791
1832	84521663 (3663, 3664)	Novel Protein sim. GBank gi 1330345 (U58755) - coded for by C. elegans cDNA yk34b1.5; coded for by C. elegans cDNA yk13h10.5; coded for by C. elegans cDNA yk46e8.5; coded for by C. elegans cDNA yk46d5.5; coded for by C. elegans cDNA yk43c2.5; coded for by C. elegans cDNA yk46e8....			264602

1833	95314184 (3665, 3668)	Novel Protein sim. GBank gi 5174413 ref NP_006026.1 pCDC4 - CDC42-binding protein kinase beta (DMPK-like)	Contains protein domain (PF00089) - Eukaryotic protein kinase domain	kinase	22278994, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 60432289, 29331827, 35686052, 29146499, 264508, 264509, 264808, 264907, 66712502, 264908, 52844045, 264909, 264512, 265008, 264591, 264593, 60433356, 21908754, 33657084, 265011, 265017, 264604, 265018, 265019, 264681, 18108351, 264683, 264288, 264685, 264768, 264687, 21908785, 21908786, 21908787, 21908789, 21908789, 29148829, 265020, 265021, 264690, 264692, 33657023, 65274620, 33657182, 27486284, 33657349, 65274791, 264634, 264635, 264556, 264557, 264558, 264559, 18108385, 58526486, 87168518, 60432113, 22279000, 22279002, 264563
1834	80582780 (3687, 3688)				264259, 264907, 264689, 22279000, 22279002
1835	94135718 (3669, 3670)			UNCLASSIFIED	22278998, 29331822, 29331826, 87168474, 264603, 21908788, 263978, 35685855, 83373044
1836	87348450 (3671, 3672)	Novel Protein sim. GBank gi 4759286 ref NP_004268.1 pUCP4 - uncoupling protein 4	Contains protein domain (PF00153) - Mitochondrial carrier proteins	transport	29331825, 264908, 265019, 264784, 264688, 21908765, 264635
1837	94234297 (3673, 3674)	Novel Protein sim. GBank gi 3334400 sp Q24574 UBPE_DROME - UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 64E (UBIQUITIN THIOLESTERASE 64E) (UBIQUITIN-SPECIFIC PROCESSING PROTEASE 64E) (DEUBIQUITINATING ENZYME 64E)	Contains protein domain (PF00443) - Ubiquitin carboxyl-terminal hydrolase family 2	ubiquitin	22278995, 29146499, 265006, 265008, 265009, 265010, 264683, 21908765, 29148827, 29148629, 265020, 265021, 265022, 65274620, 18108370, 18108374, 264556, 18108385
1838	94324369 (3675, 3676)	Novel Protein sim. GBank gi 1362599 ref A56154 - Abi substrate ena (enabled) - fruit fly (Drosophila melanogaster)	Contains protein domain (PF00568) - WH1 domain		29331822, 265017, 264760, 265020, 83373044
1839	87456508 (3677, 3678)	Novel Protein sim. GBank gi 2117310 emb CAB09116.1 - (Z95620) hypothetical protein [Schizosaccharomyces pombe]		UNCLASSIFIED	60433438, 264601, 21908765, 21908768, 265021, 33657109, 264556
1840	87391708 (3679, 3680)	Novel Protein sim. GBank gi 127560 sp P23249 MOV10_MOUSE - PROTEIN MOV-10		UNCLASSIFIED	264693
1841	85818445 (3681, 3682)	Novel Protein sim. GBank gi 4572484 gb AAD23834.1 AF12365 - (AF123653) FEZ1 [Homo sapiens]			56182575, 29331824, 29331826, 60433356, 264784, 264288, 33657023, 283987, 18108370, 18108374, 264631, 264555, 264556, 264639

1842	90982645 (3683, 3684)	Novel Protein sim. GBank gij1326268 (U56728) - C54H2.1 gene product [Caenorhabditis elegans]		UNCLASSIFIED	65274572, 29331822, 29331824, 29331825, 66714117, 29331826, 29331827, 264907, 264909, 52644045, 56182435, 264510, 265006, 265007, 265008, 264910, 60433356, 264757, 60433438, 55812038, 265017, 18108354, 264686, 264769, 33657023, 264693, 18108364, 33657109, 18108368, 264628, 55810784, 56182323, 18108384, 264563, 264584
1843	95292682 (3685, 3686)			UNCLASSIFIED	264488, 56182435, 264769, 29331828, 29331828, 264511, 265006, 265007, 264910, 264631, 264509, 264690, 264636, 264584, 264691, 60432229, 60432049, 264259, 264629, 33657023, 264486, 264909, 264587, 264595, 264768
1844	87444784 (3687, 3688)	Novel Protein sim. GBank gij2498887 (spjQ09232) YQ22_CAEEL - HYPOTHETICAL 32.0 KD PROTEIN C09F5.2 IN CHROMOSOME III		UNCLASSIFIED	264908, 265022, 33657023, 87168518, 22279002
1845	95086673 (3689, 3690)	Novel Protein sim. GBank gij175494 (spjQ09819) YAC5_SCHPO - HYPOTHETICAL 45.0 KD PROTEIN C16C9.05 IN CHROMOSOME I	Contains protein domain (PF00628) - PHD-finger	transcript factor	264259, 29331824, 264907, 264908, 66715202, 264510, 265007, 265008, 55812038, 265018, 21908765, 52644150, 33657109, 264555, 264556, 264557, 56182323, 18108382, 83373044, 18108385, 264584
1846	84287872 (3691, 3692)	Novel Protein sim. GBank gij388108 (embjCAA21739) - (AL032857) similar to EGF-like domain; cDNA EST yk299a12.3 comes from this gene; cDNA EST EMBL:D35398 comes from this gene; cDNA EST yk331h6.5 comes from this gene; cDNA EST yk299a12.5 comes from this gene; cDNA EST yk467g8.5 c...	Contains protein domain (PF00008) - EGF-like domain	UNCLASSIFIED	264905, 264908
1847	87821497 (3693, 3694)	Novel Protein sim. GBank gij5059323 (gbjAAD3867.1) AF151522 - (AF151522) hairy and enhancer of split related-1 [Homo sapiens]	Contains protein domain (PF00010) - Helix-loop-helix DNA-binding domain	transcript factor	22278997, 264259, 29331824, 264908, 18108351, 263974, 22279002
1848	86789360 (3695, 3696)	Novel Protein sim. GBank gij3701834 (embjCAB52191.1) - (AJ245417) G5b protein [Homo sapiens]	lm7		29331825, 29331826, 29331827, 265017, 264683, 264288, 264768, 264768, 21908767, 21908768, 264692, 22279002
1849	84287874 (3687, 3698)	Novel Protein sim. GBank gij450365 (reijNP_001989.1) pFBLN - fibulin 2 precursor	Contains protein domain (PF00008) - EGF-like domain	ATPase associated	56182575, 265018
1850	86689650 (3699, 3700)	Novel Protein sim. GBank gij4589582 (gbjBAA76813.1) - (AB023188) KIAA0969 protein [Homo sapiens]	Contains protein domain (PF00169) - strud		60432049, 264908
1851	95418789 (3701, 3702)	Novel Protein sim. GBank gij220637 (gbjBAA01471) - (D10627) zinc finger protein [Mus musculus]	Contains protein domain (PF00098) - Zinc finger, C2H2 type	dna_ma_bind	29331824, 35696052, 264910, 60433438, 264688, 35695917, 265020, 52644150, 65274620, 52644332

1852	95413170 (3703, 3704)	Novel Protein sim. GBank gi 5174829 ref NP_006090.1 pPIAS - protein inhibitor of activated STAT3		UNCLASSIFIED	56182575, 35686288, 22278996, 22278997, 22278999, 284490, 60432049, 284259, 29331822, 29331824, 29331825, 29331826, 60432289, 29331827, 35696052, 52844045, 285007, 284910, 60432229, 60433358, 60433438, 55812038, 65274444, 285018, 265019, 18108351, 284448, 264688, 264687, 21908765, 21908767, 21908769, 285021, 265022, 52844150, 284683, 33857109, 18108370, 18108374, 55811576, 35695855, 56182323, 60432113, 22279002, 264563 264687, 264788, 52844507, 284789, 21908765, 21906787, 21906788, 22278995, 56994075, 22278999, 52844150, 284259, 264692, 29331822, 29331824, 52845129, 29331827, 33856970, 33657349, 35895763, 284508, 264908, 284628, 264907, 284628, 284909, 35698423, 35895855, 284510, 285008, 284511, 284512, 264630, 265009, 284631, 284910, 284634, 284635, 284637, 284593, 264638, 264639, 33657402, 18108385, 52846317, 52844288, 87188518, 87188559, 284602, 285017, 22279000, 285018, 284780, 284782, 284682, 284448, 264784, 284684, 284587, 284288, 284389, 264788 264592
1853	91222287 (3705, 3706)	Novel Protein sim. GBank gi 854065 emb CAA58337 - (X83413) U88 (Human herpesvirus 6)		UNCLASSIFIED	
1854	86038152 (3707, 3708)	Novel Protein sim. GBank gi 2072864 (U93569) - putative p150 [Homo sapiens]		nuclease	
1855	91221459 (3709, 3710)	Novel Protein sim. GBank gi 4539520 emb CAB39984.1 - (AL035424) dA22D12.1 (novel protein similar to Drosophila Kelch (Ring Canal protein, KEL) and a heterogeneous set of other types of proteins) [Homo sapiens]	Contains protein domain (PF01344) - Kelch motif	nucl_rept	18108392, 52846365, 65274572, 56182575, 22278994, 56994075, 22278998, 35696288, 22278999, 284259, 60432049, 284908, 264510, 265007, 265008, 265009, 264595, 21908754, 87168474, 285011, 87168559, 264681, 264288, 264788, 18108359, 21908764, 21906788, 29148627, 285020, 265021, 265022, 52844150, 33657023, 33857109, 18108372, 18108374, 18108376, 35698423, 264631, 264638, 18108381, 264482
1856	94231871 (3711, 3712)	Novel Protein sim. GBank gi 3954878 emb CAA06945 - (AJ006278) acetylglucosaminyltransferase-like protein [Mus musculus]		UNCLASSIFIED	56994075, 264259, 29331828, 264511, 284910, 284758, 284693, 264637, 18108381, 83373044
1857	94324455 (3713, 3714)	Novel Protein sim. GBank gi 4322670 gb AAD16120 - (AF084508) denlin phosphoryn [Homo sapiens]		ATPase_associated	22278999, 284259, 264908, 60170831, 264448, 264688, 285020, 265022, 33657109, 60170394, 83373044

1858	87628311 (3715, 3716)	Novel Protein sim. GBank gi 4981803 gb AAD38415.1 AE00178 - (AE001788) ribosomal protein S15 [Thermotoga maritima]	Contains protein domain (PF00312) - Ribosomal protein S15	ribosomal prot	264757
1859	84407464 (3717, 3718)	Novel Protein sim. GBank gi 4240317 dbj BAA74837.1 - (AB020721) KIAA0814 protein [Homo sapiens]			22278998, 29331824, 265007, 33109954, 265019, 264369, 21808788, 29148784, 27488261, 52644332, 22278002 265019
1860	17828308 (3719, 3720)	Novel Protein sim. GBank gi 4009522 (AF099731) - connexin 31.1 [Homo sapiens]			
1861	88086370 (3721, 3722)	Novel Protein sim. GBank gi 2143637 pir j B4505 - calcium- dependent actin-binding protein - rat	Contains protein domain (PF00285) - Citrate synthase		264887, 284259, 29331822, 29331824, 29331825, 265007, 265009, 284591, 33109954, 265010, 265019, 264369, 264288, 264686, 264691, 264693, 27488264, 18108370, 18108374, 263977, 55811576, 58182323, 264639, 22279000, 22279002, 264482
1862	87373923 (3723, 3724)	Novel Protein sim. GBank gi 125493 sp P07313 KMLC_RABIT - MYOSIN LIGHT CHAIN KINASE, SKELETAL MUSCLE (MLCK)	Contains protein domain (PF00089) - Eukaryotic protein kinase domain	kinase	35898288, 264259, 87168474, 264369, 21908768, 264558, 264583
1863	85775037 (3725, 3726)	Novel Protein sim. GBank gi 3820809 emb CAA08299 - (AJ010842) Dof protein [Drosophila melanogaster]		UNCLASSIFIED	264601, 284768, 29148827, 29148828, 284692, 264629, 264635
1864	85547832 (3727, 3728)	Novel Protein sim. GBank gi 4322263 gb AAD15985 - (AF077738) metalloproteinase CPX-1 [Mus musculus]	Contains protein domain (PF00754) - F5/8 type C domain	synthase	22278998, 284259, 284907, 265018, 18108370, 264634, 264635, 284555, 264558, 264638, 18108387
1865	87740827 (3728, 3730)	Novel Protein sim. GBank gi 2485727 sp Q93073 Y258_HUMAN - HYPOTHETICAL PROTEIN KIAA0258			22278998, 264490, 29331822, 66714117, 66712502, 265008, 265007, 265008, 265009, 264591, 60433438, 285010, 265019, 264760, 264448, 284768, 29148827, 29148829, 265020, 265022, 18108385, 60432113
1866	87266816 (3731, 3732)	Novel Protein sim. GBank gi 5282617 emb CAB45748.1 - (AL080157) hypothetical protein [Homo sapiens]		kinase	18108374, 264769, 18108377, 21908765, 21808766, 35898423, 58182575, 21908769, 29148828, 35898288, 35895917, 265021, 264510, 284511, 284512, 264534, 284535, 60170831, 52644150, 264555, 264691, 264259, 284556, 264692, 264557, 33657023, 60433358, 29331822, 264559, 264595, 29331824, 18108385, 21908754, 33657182, 29331827, 35898052, 33658970, 87168518, 265017, 60431602, 22279000, 264508, 284509, 18108351, 264907, 264682, 264587, 18108372, 264765, 264488
1867	84579159 (3733, 3734)	Novel Protein sim. GBank gi 3859830 (AF078096) - forkhead/winged helix-like transcription factor 7 [Homo sapiens]		UNCLASSIFIED	264094

1868	87357459 (3735, 3738)	Novel Protein sim. GBank gl 3881525 emb CAA93884 - (Z70038) cDNA EST EMBL:D25279 comes from this gene; cDNA EST EMBL:D35254 comes from this gene; cDNA EST yk224b3.5 comes from this gene; cDNA EST yk357110.5 comes from this gene [Caenorhabditis elegans]	nuclease	264489, 22278997, 22278999, 29331825, 29331828, 265008, 265009, 33857402, 87188474, 18108351, 21906785, 21906788, 21906789, 265020, 265021, 60170815, 27486284, 284828, 18108374, 264831, 18108385, 87168518, 22279000, 22279002, 264568, 264567
1869	86977292 (3737, 3738)	Novel Protein sim. GBank gl 4826772 ref NP_004961.1 p GFA - Insulin-like growth factor binding protein, acid labile subunit	glycoprotein	264508, 264509, 264808, 264808, 264809, 264910, 264591, 264600, 18108351, 264683, 264768, 264789, 35695855, 264634, 264558, 264639, 18108385, 264593, 264488
1870	95349488 (3739, 3740)	Novel Protein sim. GBank gl 1889859 emb CAB08722 - (Z86099) very large tegument protein [human herpesvirus 2]	UNCLASSIFIED	29331824, 60424269, 265007, 265008, 21908754, 285017, 265018, 265019, 264288, 264766, 264686, 264688, 21906788, 21908769, 35695917, 80170815, 264692, 18108388, 35695783, 35696423, 65274791, 264638, 264639, 58528488
1871	80234464 (3741, 3742)		UNCLASSIFIED	264509, 264905, 264595, 264768, 264635, 264636, 264583, 264486
1872	80235355 (3743, 3744)	Novel Protein sim. GBank gl 2460318 (AF022147) - uterus-ovary specific putative transmembrane protein [Rattus norvegicus]	protease	264510, 264594, 264565
1873	80213890 (3745, 3746)			264509, 264512, 265009, 265011, 18108351, 264887, 264691, 18108370, 18108374, 264635
1874	95351138 (3747, 3748)		UNCLASSIFIED	264488, 35695917, 264259, 264905, 264907, 264908, 264909, 263978, 264511, 264835, 264638, 264637, 264838, 33837402, 264558, 18108385, 264600, 264604, 264764, 264587, 264788
1875	87330518 (3749, 3750)	Novel Protein sim. GBank gl 4569520 dbj BAA76782.1 - (AB023155) KIAA0938 protein [Homo sapiens]	UNCLASSIFIED	35696286, 264828, 264592, 264557, 264558
1876	87112950 (3751, 3752)	Novel Protein sim. GBank gl 263810 bbs 122920 - collagen alpha chain [Rattus ptychocheilus=tube worms, Peptide, 1027 aa]	UNCLASSIFIED	264259, 29331822, 60432289, 264908, 264909, 264604, 264784, 264288, 264769, 18108376, 264558, 264558, 264559, 18108385
1877	87315208 (3753, 3754)	Novel Protein sim. GBank gl 3983356 gb AAC83924.1 - (AF102545) riboflavin binding protein precursor [Scaphiopus couchii]	UNCLASSIFIED	264767, 264688, 264768, 263978, 264693, 264639, 265010, 264583, 264905, 264806, 264907

1878	95351056 (3755, 3756)	Novel Protein sim. GBank gij4510345jigbiAAD21434.11 - (AC006921) unknown protein (Arabidopsis thaliana)	Contains protein domain (PF01428) - AN1-like Zinc finger	ubiquitin	264589, 264488, 35696286, 56994075, 264259, 29331822, 29331824, 29331825, 3569052, 29331828, 29148498, 264508, 264509, 264905, 264908, 264907, 264908, 264909, 264510, 264511, 264512, 265009, 264910, 33657402, 264594, 264758, 55812038, 265011, 264602, 264760, 18108351, 264762, 264681, 264682, 264764, 264389, 264288, 264766, 264687, 264768, 264789, 21906768, 35695917, 265021, 60170615, 33657023, 264692, 264693, 33657109, 27486285, 264628, 18108370, 264629, 18108374, 35698423, 264634, 264635, 264555, 264638, 264639, 83373044, 18108385, 56526486, 87188518, 264583, 264594, 264586, 264488, 264587
1879	95310883 (3757, 3758)	Novel Protein sim. GBank gij4928643jigbiAAD34082.1jAF15184 - (AF151845) CGI-87 protein (Homo sapiens)		UNCLASSIFIED	264905, 264907, 264908, 265007, 264565, 264566
1880	91012978 (3759, 3760)	Novel Protein sim. GBank gij1550785jemb[CAA69283] - (Y08026) Immune associated protein 38 [Mus musculus]		UNCLASSIFIED	264766, 264691, 264692, 83373044
1881	80214949 (3761, 3762)	Novel Protein sim. GBank gij93144jipj[B40505] - hypothetical protein - suid herpesvirus 1 (strain Indiana-Funkhouser or Becker)		UNCLASSIFIED	264509, 264905, 264908, 264909, 264910, 264762, 264687, 33657023, 264632
1882	86582450 (3763, 3764)	Novel Protein sim. GBank gij2384956 (AF022985) - No definition line found [Caenorhabditis elegans]			264908, 21906768, 18108370, 263374, 87188518
1883	84216817 (3765, 3766)	Novel Protein sim. GBank gij1351218jipjP47228jTES2_MOUSE - TESTIN 2 (TES2) [CONTAINS: TESTIN 1 (TES1)]	Contains protein domain (PF00412) - LIM domain containing proteins	homeobox	264908, 264910, 87188559, 21906766, 264638

1884	95310885 (3767, 3768)	Novel Protein sim. GBank gi 4929843 gb AAD34082.1 AF15184 - (AF151845) CGI-87 protein [Homo sapiens]		UNCLASSIFIED	264488, 18108394, 58161686, 22278998, 60432049, 284259, 29331822, 29331824, 29331825, 68714117, 29331827, 35696052, 284508, 284509, 284905, 284908, 284907, 284908, 284909, 58182435, 284510, 284511, 265007, 284512, 285009, 284910, 284592, 284593, 33657402, 55812038, 284758, 55811388, 285010, 285011, 87168559, 284600, 285017, 284604, 285019, 284605, 284760, 55811150, 284761, 284682, 284763, 284683, 284764, 284288, 284369, 284768, 284688, 284768, 284769, 29148764, 35695917, 284890, 284891, 33657023, 284892, 284693, 33657109, 18108370, 284628, 60431528, 284629, 283973, 18108374, 55810764, 55811576, 35696423, 35695955, 284631, 284634, 60431850, 284637, 284638, 56182323, 284639, 18108382, 83373044, 18108385, 60432113, 22278002, 284563, 284564, 284565, 284568, 284488, 284567
1885	87644280 (3768, 3770)	Novel Protein sim. GBank gi 2507155 sp P37370 VRP1_YEAST - VERPROLIN		UNCLASSIFIED	58182575, 284259, 284905, 284909, 285008, 284598, 284768, 285020, 284628, 60431528, 284634, 58528488, 284080, 284563
1886	86674082 (3771, 3772)	Novel Protein sim. GBank gi 2854158 gb AAC02577.1 - (AF045841) No definition line found [Caenorhabditis elegans]			22278998, 22278999, 60432049, 284910, 285018, 284766, 21906768, 29148629, 284690, 284693, 284828, 284555, 284488
1887	94139139 (3773, 3774)	Novel Protein sim. GBank gi 5174421 ref NP_008023.1 pCPNE - copine VI (neuronal)	Contains protein domain (PF00188) - C2 domain	ATPase_associated	29331822, 29331824, 29331825, 29331828, 29331827, 284908, 285007, 284681, 284768, 29148627, 284693, 18108364, 35898423, 63274791, 35895855, 284632, 56182323, 284639, 284583
1888	87822804 (3775, 3776)	Novel Protein sim. GBank gi 3319931 emb CAB10841 - (Z88046) dJ1409.2 (Melanoma-Associated Antigen MAGE LIKE) [Homo sapiens]	Contains protein domain (PF01454) - MAGE family		283978
1889	91255783 (3777, 3778)	Novel Protein sim. GBank gi 1083308 pir [A56559 - enhancer-trap-locus-1 protein - mouse (fragment)]	Contains protein domain (PF00176) - SNF2 and others N-terminal domain	helicase	284906, 55812038, 284758, 285010, 285018, 285019, 18108351, 284288, 21908765, 21908768, 21908769, 285021, 33657023, 33657109, 56182323, 83373044, 18108385, 22278000, 22279002
1890	87626705 (3779, 3780)	Novel Protein sim. GBank gi 4240185 dbj BAA74876.1 - (AB020680) KIAA0853 protein [Homo sapiens]		UNCLASSIFIED	18108398, 29147620, 284907, 285009, 284600, 285018, 18108351, 284288, 284689, 21908765, 21908768, 21908769, 284691, 284692, 284693, 284628, 18108370, 284636, 284558, 284404

1891	87013895 (3781, 3782)				UNCLASSIFIED	264686, 264768, 264687, 264692, 264693, 29331822, 29331824, 264508, 264805, 264906, 18108370, 264628, 264907, 264908, 264909, 18108379, 265007, 265008, 264910, 264632, 264591, 264639, 264596, 18108384, 265010, 265011, 264601, 264605, 264583, 264389
1892	87642825 (3783, 3784)	Novel Protein sim. GBank gi 5688535 dbj BAA83051.1 - (AB028022) KIAA1089 protein [Homo sapiens]		Contains protein domain (PF01412) - Putative GTP-ase activating protein for Arf	struct	22278995, 264509, 87168559, 18108351, 264446, 264682, 265020, 264693, 18108374, 22279000
1893	88533828 (3785, 3786)				laminin	264569, 65274572, 22278997, 22278999, 264259, 29331822, 29331824, 68714117, 29331828, 264806, 265008, 265008, 265009, 264592, 265018, 264681, 264448, 264683, 18108354, 264389, 264684, 264685, 264768, 264687, 264689, 21908788, 265020, 265022, 60170815, 52644150, 264890, 264691, 264692, 33657023, 264693, 33657109, 264628, 18108374, 35858555, 264630, 264632, 264634, 264557, 264558, 60170394, 18108381, 18108385, 22279000
1894	86989120 (3787, 3788)					264508, 264905, 264906, 264907, 264594, 264684, 264680, 264692, 264630, 264635, 264636, 264639, 264583
1895	87631891 (3789, 3790)	Novel Protein sim. GBank gi 5282574 emb CAB45729.1 - (AL080133) hypothetical protein [Homo sapiens]		Contains protein domain (PF00435) - Spectrin repeat		58182575, 264259, 60432289, 29331828, 284107, 264905, 264908, 264910, 60170831, 264758, 265010, 265018, 264448, 264288, 264768, 33657109, 264628, 55810764, 18108379, 264634, 58182323, 58526488
1896	85673555 (3791, 3792)			Contains protein domain (PF00627) - UBA domain	UNCLASSIFIED	264907, 265008, 264682, 264686, 21908768, 264629, 264631, 264634, 264555
1897	80565569 (3793, 3794)	Novel Protein sim. GBank gi 72883 sp P39193 ALU6 SP WARNING ENTRY IIII			cadherin	264259
1898	87617637 (3795, 3796)	Novel Protein sim. GBank gi 127560 sp P23248 MV10_MOUSE - PROTEIN MOV-10			helicase	22278996, 22278998, 22278999, 29331824, 29331825, 60432289, 29331827, 35696052, 29331828, 265008, 265019, 264681, 264682, 264448, 264369, 52644229, 21908765, 21908768, 21908768, 21906769, 60170815, 55810764, 22279000
1899	86673087 (3797, 3798)	Novel Protein sim. GBank gi 2909819 (AF031548) - erythrocyte membrane glycoprotein Rh50 [Homo sapiens]		Contains protein domain (PF00809) - Ammonium Transporter Family	glycoprotein	264259, 264508, 264909, 60432229, 264769, 21908765, 21906769
1900	87641858 (3799, 3800)	Novel Protein sim. GBank gi 4102881 (AF017250) - vitellogenin precursor [Oreochromis aureus]			UNCLASSIFIED	264683

1801	95186847 (3801, 3802)	Novel Protein sim. GBank glij585859[spjP38378]S61A_RAT PROTEIN TRANSPORT PROTEIN SEC61 ALPHA SUBUNIT	Contains protein domain (PF00242) - DNA polymerase (viral) N-terminal domain	transport	264488, 52844507, 52845156, 18108396, 52846365, 52846842, 18108397, 56182575, 22278994, 22278995, 56994075, 35896286, 22278997, 22278998, 264490, 60432049, 264259, 29331822, 52845080, 29331824, 29331825, 68714117, 29331826, 60432289, 29331827, 29331828, 35898032, 33658970, 29146498, 284906, 264907, 29331830, 284908, 52844045, 284909, 284112, 285008, 284512, 285008, 284910, 285009, 60170831, 60432228, 60433356, 33657402, 60433438, 55812038, 264758, 33109854, 21908754, 33857084, 52844298, 87168474, 285010, 285011, 87168559, 285017, 265018, 285019, 18108351, 284448, 284288, 264886, 52844229, 21908765, 21806786, 21908767, 21908769, 55811957, 35895917, 285020, 285021, 52844150, 18108382, 33657023, 264893, 263987, 33657109, 33657182, 27488264, 33857349, 35895763, 18108370, 18108378, 55811578, 35898423, 35895855, 60431850, 284636, 283981, 52844332, 60170394, 83373044, 18108385, 87168518, 60432113, 284584 264107, 263978
1802	80202013 (3803, 3804)	Novel Protein sim. GBank glij426613[gbjAAD20451] - (AF098798) SLM-1 [Mus musculus]		dna_rna_bind	
1803	87778554 (3805, 3806)	Novel Protein sim. GBank glij3747107 (AF095741) - unknown [Rattus norvegicus]		UNCLASSIFIED	264259, 29331825, 29331827, 284508, 264907, 285008, 60170831, 60433356, 60433438, 264759, 21908754, 264448, 264288, 285021, 285022, 33857023, 264693, 55811578, 284555, 284558, 22279000
1804	80434213 (3807, 3808)	Novel Protein sim. GBank glij1352911[spjP47147]YJ80_YEAST - HYPOTHETICAL 80.2 KD PROTEIN IN CPA2-NNF1 INTERGENIC REGION		stuctl	284509, 284905, 284908, 284907, 284908, 285007, 264910, 284686, 264788, 264687, 264789, 264693, 284628, 18108374, 284634, 284638, 264637, 284585
1805	95351140 (3809, 3810)	Novel Protein sim. GBank glij3043714[dbj]BAA25521 - (AB011167) KIAA0595 protein [Homo sapiens]	Contains protein domain (PF00293) - Bacterial muT protein		284488, 264768, 284789, 284689, 29148629, 35895917, 35896286, 264259, 284692, 18108362, 33857023, 29331824, 33857109, 29148499, 284508, 284509, 284905, 284906, 284907, 68712502, 284908, 284909, 35898423, 35895855, 284510, 284511, 284512, 264910, 284634, 284635, 284637, 284638, 33657402, 284758, 85658542, 284802, 264760, 264781, 264482, 284583, 264782, 284483, 284784, 284566, 284288, 264768
1806	12763822 (3811, 3812)			UNCLASSIFIED	284637

1907	85351144 (3813, 3814)	Novel Protein sim. GBank gi 4929585 gb AAD34053.1 AF15181 - (AF15181) CGI-58 protein [Homo sapiens]	Contains protein domain (PF00561) - alpha/beta hydrolase fold	-hydrolase	65274572, 22278998, 35698286, 22278998, 22278998, 264259, 60432049, 29331822, 29331825, 60424269, 29331828, 35696052, 264593, 60433356, 21908754, 55811386, 85858542, 87188559, 265018, 264681, 264682, 264684, 264288, 21908765, 21908768, 21908768, 265020, 265022, 264680, 52644150, 264692, 33657023, 264693, 33657109, 35695855, 264636, 264638, 60432113
1908	95313641 (3815, 3816)	Novel Protein sim. GBank gi 3888770 (AF109808) - NG22 [Mus musculus]		UNCLASSIFIED	284488, 65274572, 56182575, 56181686, 22278998, 22278998, 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 60432289, 29331826, 29331827, 29331828, 29148499, 264805, 264806, 264907, 264908, 66712502, 264809, 56182435, 264510, 284511, 265008, 264910, 60170831, 284592, 60433356, 33657402, 264594, 60433438, 264598, 55812038, 33109954, 52646317, 265011, 265017, 264604, 265018, 265019, 264805, 55811150, 264681, 264448, 284288, 264688, 264688, 264769, 21908765, 21908766, 21908767, 21908768, 21908769, 55811957, 29148628, 35695917, 265020, 265022, 264691, 264692, 18108364, 65274620, 33657109, 33657349, 35695763, 18108374, 263978, 55810764, 55811576, 35696423, 65274791, 264631, 264632, 264556, 264557, 60170394, 56182323, 83373044, 18108385, 60432113, 22278000, 22278002, 264568, 264488
1909	85514505 (3817, 3818)	Novel Protein sim. GBank gi 2224653 db BAA20813 - (AB002354) KIAA0358 [Homo sapiens]		UNCLASSIFIED	264259, 264508, 264805, 264906, 264907, 264908, 264511, 264910, 264593, 264758, 264764, 264766, 18108370, 264634, 264637, 264488
1910	84216821 (3819, 3820)	Novel Protein sim. GBank gi 1351218 sp P47228 TES2_MOUSE - TESTIN 2 (TES2) [CONTAINS: TESTIN 1 (TES1)]	Contains protein domain (PF00412) - LIM domain containing proteins	homeobox	35698286, 22278998, 22278999, 35698052, 264509, 264805, 264908, 264907, 264908, 264909, 264511, 265007, 264512, 264910, 264758, 265011, 264601, 264602, 264604, 264605, 264761, 264764, 264288, 264768, 264768, 264687, 264769, 35695917, 265021, 52644150, 264692, 264628, 18108370, 264629, 18108372, 18108374, 35695855, 264631, 264634, 264635, 264636, 264637, 264638, 18108385, 264565, 264568, 264488

1811	91725345 (3821, 3822)	Novel Protein sim. GBank gi4809339 gb AAD30184.1 AC006530 hypothesized protein [Homo sapiens]	Contains protein domain (PF01119) - DNA mismatch repair protein	nuclease	18108394, 56182575, 56182181, 29331826, 29331827, 33656970, 264908, 265007, 264591, 55812038, 87188559, 264448, 264369, 21908765, 21908768, 265022, 264691, 264693, 18108385, 55811576, 264556, 18108385, 18108388
1812	95413518 (3823, 3824)	Novel Protein sim. GBank gi5689439 dbj BAA83003.1 - (AB028974) KIAA1051 protein [Homo sapiens]	Contains protein domain (PF00098) - Zinc finger, CCHC class	UNCLASSIFIED	18108397, 56182575, 56181686, 22278994, 22278995, 56994075, 22278996, 22278999, 264259, 29331822, 29331824, 56182181, 29331825, 66714117, 35896052, 264905, 264906, 264907, 264908, 52844045, 56182435, 265007, 265008, 264910, 265009, 264591, 264596, 85274444, 55811386, 87188474, 265011, 87188559, 265018, 265019, 264760, 18108351, 264681, 264369, 264684, 264288, 264688, 264768, 21908765, 21908766, 21908767, 21908769, 21908769, 55811957, 265020, 265021, 265022, 80170815, 264692, 33657023, 264693, 18108376, 55811576, 35896423, 85274791, 264637, 56182323, 83373044, 56526466, 22278002, 264563, 264568
1813	95305546 (3825, 3826)	Novel Protein sim. GBank gi5032245 ref NP_005665.1 pZNF2 - zinc finger protein (C2H2) homologous to mouse MOK-2			56182575, 22278994, 22278995, 56994075, 22278996, 22278998, 22278999, 29331826, 29331827, 265008, 55812038, 265010, 265017, 265018, 265019, 264681, 18108351, 264683, 264764, 264369, 264288, 264685, 264686, 264769, 21908765, 21908766, 21908768, 21908769, 55811957, 265020, 265022, 264691, 55811576, 264634, 264635, 264638, 56182323, 83373044, 18108385
1814	83423982 (3827, 3828)	Novel Protein sim. GBank gi4589604 dbj BAA78624.1 - (AB023197) KIAA0980 protein [Homo sapiens]	Contains protein domain (PF00036) - EF hand	struct	56182575, 29331824, 35896052, 264908, 264908, 264626, 264909, 264592, 264758, 87188559, 18108351, 18108354, 264684, 264686, 33657023, 264693, 264628, 264631, 264632, 264634, 264635, 264639
1815	95340459 (3829, 3830)	Novel Protein sim. GBank gi5689415 dbj BAA82991.1 - (AB028982) KIAA1039 protein [Homo sapiens]		UNCLASSIFIED	264259, 29331824, 29331826, 29331827, 264508, 264909, 265009, 265017, 265019, 264768, 264769, 264689, 264628, 264635, 264637, 264639, 83373044, 264565
1816	79640761 (3831, 3832)				264693, 264639

1917	87821880 (3833, 3834)	Novel Protein sim. GBank gij5689391[dbj BAA02978.1] - (AB028950) KIAA1027 protein [Homo sapiens]		struct	284769, 284689, 21908765, 21908768, 22278998, 284259, 284691, 284693, 29331824, 29331825, 29331826, 29331828, 284805, 284908, 284628, 284907, 284908, 284909, 284510, 284630, 284910, 284634, 284635, 284638, 284637, 284638, 283981, 284639, 284758, 18108385, 21908754, 285011, 284604, 284583, 18108351, 284762, 284763, 284588, 284764, 284768
1918	85302785 (3835, 3836)	Novel Protein sim. GBank gij5281517[gb AAD41524.1 AF154831] PV-1 [Rattus norvegicus]		struct	284488, 18108392, 18108357, 21908765, 21908767, 21908768, 58182575, 21908769, 22278994, 35696286, 35695917, 22278998, 22278997, 265021, 265022, 284534, 284690, 284691, 284692, 33657023, 284693, 29331824, 29331825, 33857109, 29331826, 52845129, 35696052, 29331828, 27486282, 27486284, 35695783, 284508, 284905, 284509, 284808, 284628, 284907, 18108370, 284908, 284629, 284909, 18108372, 18108374, 263978, 35696423, 35695855, 284510, 284511, 285006, 285007, 284512, 265008, 284831, 285009, 284910, 284634, 284635, 284555, 284636, 284556, 284637, 284557, 284593, 284638, 284594, 60170394, 284595, 284559, 284598, 8337304, 284758, 52846317, 18108385, 52844288, 58528488, 87168518, 265010, 265011, 87168559, 284600, 284601, 284602, 285017, 284603, 284604, 285018, 284605, 284780, 284781, 284482, 284584, 18108351, 284762, 284692, 284585, 284448, 284764, 284586, 284488, 284587, 284389, 284288, 284768, 284487, 284685
1919	94143847 (3837, 3838)	Novel Protein sim. GBank gij3878584[emb CAB01237] - (Z77667) cDNA EST EMBL:CO8125 comes from this gene; cDNA EST EMBL:CO9753 comes from this gene [Caenorhabditis elegans]		oxidase	22278997, 29331822, 265007, 60170831, 80432229, 80433438, 284448, 284692, 284288, 55811957, 33657023, 33657109, 85274791, 58182323, 22279002
1920	91228953 (3839, 3840)	Novel Protein sim. GBank gij1809231 (AC000115) - coded for by human cDNAs R76043 (NID:9850725), R65857 (NID:9838495) and H12888 (NID:9877688) [Homo sapiens]		UNCLASSIFIED	284510, 284511, 284512, 284586
1921	79555226 (3841, 3842)	Novel Protein sim. GBank gij4580987[gb AAD24571.1 AF121081] CAMP inducible 2 protein [Mus musculus]		UNCLASSIFIED	284693

1922	87641863 (3843, 3844)	Novel Protein sim. GBank gi 36593sp P02845 VT2_CHICK - VITELLOGENIN II PRECURSOR (MAJOR VITELLOGENIN) [CONTAINS: LIPOVITELLIN I (LVI); PHOSVITIN (PV); LIPOVITELLIN II (LVI)]; YGP40]		UNCLASSIFIED	264686, 264688, 264490, 18108370, 264909, 18108374, 265008, 264537, 264584, 18108351
1923	94323589 (3845, 3846)	Novel Protein sim. GBank gi 119110 sp P03211 EBN1_EBV - EBNA-1 NUCLEAR PROTEIN		UNCLASSIFIED	264488, 264489, 22278995, 264094, 264258, 35898052, 264509, 264905, 264906, 264907, 264908, 66712502, 264909, 264511, 264512, 265007, 264910, 265008, 264591, 264592, 264593, 264594, 264595, 264758, 264598, 264759, 265011, 265017, 265018, 265019, 5581150, 264681, 264762, 264448, 264784, 264288, 264389, 264766, 264787, 264888, 264687, 264788, 264789, 56181582, 264689, 21908766, 264691, 33657023, 264693, 65274620, 33857109, 18108370, 264628, 264629, 35888423, 264630, 264631, 264632, 264634, 264555, 264636, 264637, 264638, 264639, 264556, 83373044, 87168518, 264583, 264584, 264585, 264586, 264587
1924	87338925 (3847, 3848)	Novel Protein sim. GBank gi 3877655 emb CAA96657 - (Z72511) possible zinc finger protein; cDNA EST EMBL:M89115 comes from this gene; cDNA EST EMBL:D71533 comes from this gene; cDNA EST EMBL:D72314 comes from this gene; cDNA EST EMBL:D75164 comes from this gene; cDNA EST EMBL:G1...	Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger)		29331826, 264906, 264908, 264595, 265017, 265018, 265019, 265021, 264691, 264693, 264637, 18108385, 264585
1925	87828338 (3849, 3850)	Novel Protein sim. GBank gi 4981903 gb AAD38415.1 AE00178 - (AE001788) ribosomal protein S15 [Thermotoga maritima]	Contains protein domain (PF00312) - Ribosomal protein S15	ribosomalprot	22278995, 22278996, 22278997, 264259, 29331824, 66714117, 29148499, 264909, 52644045, 265008, 265009, 264758, 265011, 265017, 264605, 264448, 264288, 264692, 33657109, 18108374, 60170394
1926	86094739 (3851, 3852)	Novel Protein sim. GBank gi 2246532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma- associated herpesvirus]		struct	264905, 264906, 264907, 264910, 264593, 265018, 264760, 264764, 264288, 264692, 264693, 263978, 264631, 264634, 264637, 264583
1927	85654857 (3853, 3854)	Novel Protein sim. GBank gi 3043632 dbj BAA25480 - (AB011126) KIAA0554 protein [Homo sapiens]			264689, 264631
1928	87799054 (3855, 3856)	Novel Protein sim. GBank gi 1685761 dbj BAA13377 - (D87433) KIAA0246 [Homo sapiens]	Contains protein domain (PF00183) - Inf Extracellular link domain	Inf	264489, 264259, 265017, 265021, 264692
1929	86997236 (3857, 3858)	Novel Protein sim. GBank gi 5001893 gb AAD37247.1 AF13432 - (AF134321) chimeric AFGP/rypsinogen-like serine protease precursor [Dissostichus mawsoni]		UNCLASSIFIED	264508, 264591, 33657402, 265017, 264768, 264632, 264556, 264639

1930	87889128 (3859, 3860)	Novel Protein sim. GBank gij1708230[sp]P52863[NBL4_MOUSE - NBL4 PROTEIN]		phosphatase	35686286, 29331828, 284905, 284907, 284908, 284909, 284511, 284910, 284758, 284801, 285017, 285018, 284605, 284780, 284784, 284788, 284688, 284789, 285022, 35686423, 284638, 60432113
1931	87797278 (3861, 3862)	Novel Protein sim. GBank gij404634 (U01840) - serine/threonine kinase [Mus musculus]	Contains protein domain (PF00089) - Eukaryotic protein kinase domain	kinase	284908, 284908, 60432229, 284758, 284784, 284288, 285020, 284692, 284634, 284637, 284684, 284691, 284635
1932	15030972 (3863, 3864)			UNCLASSIFIED	284595
1933	11813688 (3865, 3866)	Novel Protein sim. GBank gij4115748[bj]BAA364941 - (AB022023) nonmuscle myosin heavy chain B [Bos taurus]		struct	56182575, 56182435, 284510, 284757, 284758, 55812038, 55811388, 285018, 55811150, 21908765, 284691, 284631, 284635, 284637
1934	84426360 (3867, 3868)			UNCLASSIFIED	284688, 285011, 284511, 284905, 18108351, 284584, 284681, 284259, 18108370, 284586, 284784, 284369, 284595
1935	87752511 (3869, 3870)				60432289, 285007, 285010, 285011, 285018, 33657109, 18108374
1936	95414338 (3871, 3872)	Novel Protein sim. GBank gij4827040[re]NP_005110.1pTRAP - thyroid hormone receptor-associated protein, 150 kDa subunit			
1937	94947141 (3873, 3874)	Novel Protein sim. GBank gij543187[pj]S37771 - ankyrin, erythrocyte - mouse	Contains protein domain (PF00023) - Ank repeat	kinase	85658542, 21906787, 35695917, 60170615, 284693, 33857109
1938	87403277 (3875, 3876)	Novel Protein sim. GBank gij4544431[gb]AAD2340.1AC00895 - (AC008955) hypothelical protein [Arabidopsis thaliana]	Contains protein domain (PF00888) - Cullin family	collagen	284488, 28146498, 284905, 284559
1939	91004978 (3877, 3878)	Novel Protein sim. GBank gij500856[db]BAA03210] - (O14168) 50kDa lectin [Bombyx mori]		UNCLASSIFIED	65274572, 22278999, 60432289, 56182435, 60433356, 285017, 21908765, 21908768, 21908768, 55811957, 27486264, 35686423, 60432113, 284584
1940	87346810 (3879, 3880)	Novel Protein sim. GBank gij1946300[emb]CAA73132] - (Y12529) hypothelical protein [Silene latifolia]	Contains protein domain (PF00580) - Leucine Rich Repeat	struct	284488, 28331822, 284448, 284683, 284288, 285020, 33857023, 284631
1941	84147177 (3881, 3882)	Novel Protein sim. GBank gij4206388 (AF060570) - rig-1 protein [Mus musculus]		UNCLASSIFIED	56994075, 22278998, 284259, 29331824, 29331827, 284905, 285008, 33857084, 285017, 285018, 284288, 284687, 21908765, 21908768, 21908767, 285020, 52844150, 27486264, 83373044, 18108387, 60432113, 22278002, 284585
1942	87841870 (3883, 3884)	Novel Protein sim. GBank gij4927204[gb]AAD33049.1JAF133911) ARL-8 interacting protein-4 [Mus musculus]		UNCLASSIFIED	284488, 18108398, 29331825, 27486281, 284509, 18108370, 18108374, 284482
1943	94325288 (3885, 3886)	Novel Protein sim. GBank gij3122952[sp]O15736]TIPD_DICDI - TIPD PROTEIN	Contains protein domain (PF00400) - WD domain, G-beta repeat	kinase	22278998, 29331822, 29331827, 35686052, 284511, 265009, 284592, 60432229, 285017, 285018, 285019, 284684, 284692, 33857109, 65274791, 284638

1844	94232958 (3887, 3888)	Novel Protein sim. GBank gi 1789570 dbj BAA13432 - (D87871) TIP120 [Rattus norvegicus]	UNCLASSIFIED	65274572, 22278994, 22278995, 22278996, 22278998, 22278999, 284259, 52645080, 29331824, 29331826, 29331827, 66712502, 58182435, 284512, 284910, 60170831, 60432228, 60433356, 33657402, 60433438, 284596, 33109954, 21908754, 87168474, 87168559, 265017, 265018, 265019, 18108351, 284369, 284686, 284768, 21906765, 21906766, 21908767, 21906769, 35895917, 265020, 265021, 60170615, 284692, 33657023, 18108370, 18108374, 35896423, 35695855, 284634, 60170394, 284639, 83373044, 18108385, 58526488, 87168518, 60432113, 22279000, 284563, 284488, 22278996, 284510, 284511, 18108351, 264683, 284488, 284567
1845	87641872 (3888, 3890)	Novel Protein sim. GBank gi 4927204 gb AAD33048.1 AF13391 - (AF13391.1) ARL-6 interacting protein-4 [Mus musculus]	UNCLASSIFIED	60432289, 29331827, 35896052, 265007, 265008, 60433356, 60433438, 284369, 58181562, 21908767, 52644150, 284693, 27489264, 284637, 87168518, 284563
1846	87443990 (3891, 3892)	Novel Protein sim. GBank gi 2498104 sp Q27969 AD50_BOVIN - ADRENAL MEDULLA 50 KD PROTEIN	UNCLASSIFIED	22278998, 284805, 264906, 264908, 264909, 284512, 284758, 284762, 284682, 284683, 284764, 264286, 284768, 284688, 21906768, 284693, 18108374, 35895855, 284635, 284637, 284638, 18108385, 22279002
1847	86438862 (3893, 3894)	Novel Protein sim. GBank gi 3914801 sp O54888 RPA2_RAT - DNA-DIRECTED RNA POLYMERASE I 135 KD POLYPEPTIDE (RNA POLYMERASE I SUBUNIT 2) (RPA135) (RNA POLYMERASE I 127 KD SUBUNIT)	mapolymerase	264909, 60170831, 284591, 284594, 235010, 285011, 284764, 284369, 284688, 284631, 284638
1848	95189174 (3895, 3896)	Novel Protein sim. GBank gi 5420387 emb CAB46678.1 - (AJ243459) proteophosphoglycan [Leishmania major]	struct	284369
1849	7640129 (3897, 3898)	Novel Protein sim. GBank gi 3876766 emb CAA93466.1 - (Z69837) predicted using GeneFinder, Similarity to E. coli hypothetical protein YCAC (SW:YCAC_ECOLI) [Caenorhabditis elegans]	UNCLASSIFIED	284488, 284768, 284688, 284689, 284259, 29331822, 33657109, 29331828, 264103, 284509, 18108370, 35895855, 284510, 265008, 285009, 33657402, 18108385, 265018, 284583, 55811150, 18108351, 284369, 284288, 18108354
1850	87788531 (3899, 3900)	Novel Protein sim. GBank gi 2626753 dbj BAA23424 - (AB008782) sulfate transporter [Arabidopsis thaliana]	Contains protein domain (PF00816) - Isochoitismalase family	58182575, 22278997, 52645080, 29331824, 29331825, 29331827, 55812038, 52848317, 265018, 265019, 284369, 21906765, 21906767, 55811957, 265020, 265021, 33657023, 284683, 35695763, 58182323, 22279002
1851	86888253 (3801, 3902)	Novel Protein sim. GBank gi 2626753 dbj BAA23424 - (AB008782) sulfate transporter [Arabidopsis thaliana]	Contains protein domain (PF00816) - Sulfate transporter family	264591, 284593, 284594, 264595, 264555, 284558, 284557, 284558, 284555
1852	87069775 (3803, 3904)	Novel Protein sim. GBank gi 429833 gb AAD34077.1 AF15184 - (AF15184) CGI-82 protein [Homo sapiens]	Contains protein domain (PF00106) - short chain dehydrogenase reductase	

1853	20470371 (3905, 3906)	Novel Protein sim. GBank gij1168715sp1P31721jC1QB_RAT - COMPLEMENT C1Q SUBCOMPONENT, B CHAIN PRECURSOR (AB020698) KIAA0891 protein [Homo sapiens]	Contains protein domain (PF00388) - C1q domain	complement	264259, 264558
1854	91226025 (3907, 3908)	Novel Protein sim. GBank gij4240271jdbjBAA74914.1j - (AB020698) KIAA0891 protein [Homo sapiens]	Contains protein domain (PF00443) - Ubiquitin carboxyl-terminal hydrolase family 2	ubiquitin	264488, 263994, 18108394, 35696286, 22278998, 29331822, 68714117, 29331826, 29331827, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 265008, 265007, 265008, 264910, 265009, 80170831, 60432229, 265011, 265017, 264603, 264604, 265019, 18108351, 264685, 264786, 264769, 35695917, 265020, 265021, 264691, 33657023, 264692, 33657109, 264628, 18108374, 35698423, 55811576, 35695855, 264630, 264635, 264638, 264555, 264558, 264638, 264557, 264639, 264558, 264559, 83373044, 18108385, 264563, 264564, 264568, 264488, 264567
1855	95308310 (3909, 3910)			UNCLASSIFIED	56182575, 56181686, 22278994, 22278999, 264259, 29331822, 56182181, 29331824, 29331825, 29331826, 29331827, 35696052, 264508, 29331830, 265008, 265009, 264591, 55812038, 87168474, 265017, 265018, 265019, 264448, 264768, 21908785, 21908768, 21906787, 55811957, 265020, 265021, 52645129, 33657109, 27488264, 33657349, 35695763, 60431528, 18108374, 55811578, 35695855, 264635, 60431850, 264639, 83373044
1856	95092121 (3911, 3912)	Novel Protein sim. GBank gij1665921jdbjBAA134071 - (D87469) Similar to D.melanogaster cadherin-related tumor suppressor [Homo sapiens]	Contains protein domain (PF00028) - Cadherin domain	cadherin	264488, 56182575, 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 29331827, 35696052, 29331828, 264508, 264906, 264908, 264909, 264511, 265007, 264910, 264591, 55812038, 265010, 265018, 18108351, 264768, 56181562, 264689, 21908768, 21908769, 265022, 264691, 264628, 18108374, 55810764, 55811578, 35695855, 264631, 264632, 264635, 264637, 264639, 80170394, 56182323, 83373044, 18108385, 22279000, 22279002, 264583, 264584

1957	94326510 (3913, 3914)	Novel Protein sim. GBank gi4589674[db][BAA78856.1] - (AB023229) KIAA1012 protein [Homo sapiens]	UNCLASSIFIED	52646842, 56182575, 22278997, 22278998, 22278999, 29331824, 66714117, 29331827, 29146498, 284593, 33657402, 33109954, 87168474, 285018, 284448, 284369, 284288, 284768, 21908768, 21908767, 21908768, 21908769, 285020, 265021, 284892, 65274620, 27488264, 33657349, 27488265, 35895855, 22279002, 284482
1958	95313902 (3915, 3916)	Novel Protein sim. GBank gi4240227[db][BAA74862.1] - (AB020676) KIAA0869 protein [Homo sapiens]	UNCLASSIFIED	22278998, 284092, 264094, 264259, 80432049, 29331824, 56182181, 66714117, 284107, 284109, 284909, 284511, 60170831, 60432228, 21908754, 265010, 21908769, 35895917, 285022, 65274620, 283987, 283976, 35898423, 284631, 284632, 284634, 284635, 18108385, 22279000, 22279002, 264593, 265019
1959	85701470 (3917, 3918)	Novel Protein sim. GBank gi2281983[emb][CAB10860] - (Z98056) hypothetical protein [Schizosaccharomyces pombe]	ubiquitin	
1960	80308608 (3919, 3920)	Novel Protein sim. GBank gi227485[db][BAA21515] - (D84159) 3-7 gene product [Homo sapiens]	struct	284905, 284906, 284907, 284908, 284909, 285008, 285007, 284910, 284595, 285017, 284604, 285018, 18108351, 284764, 284369, 284768, 284768, 21908765, 18108368, 284629, 18108378, 284635, 284636, 284637, 284638, 284486
1961	16292607 (3921, 3922)			264635
1962	91008385 (3923, 3924)		UNCLASSIFIED	65274572, 264592, 264593, 265019, 284691
1963	90936017 (3925, 3926)	Novel Protein sim. GBank gi3721653[db][BAA33581] - (AB012933) acyl-CoA synthetase 5 [Rattus norvegicus]	eph	65274572, 18108398, 35896286, 29331825, 60432289, 29331827, 284828, 265006, 285009, 60433358, 60433438, 21908754, 285020, 265021, 33657023, 33657109, 27486285, 35895855, 284555
1964	94317605 (3927, 3928)	Novel Protein sim. GBank gi5262639[emb][CAB45757.1] - (AL080169) hypothetical protein [Homo sapiens]	cadherin	284486, 284092, 284259, 264509, 284905, 284808, 284907, 284908, 284909, 284510, 284511, 265007, 265009, 284910, 284592, 284593, 284594, 284595, 284758, 284600, 284603, 284604, 284605, 284760, 284762, 284448, 284764, 284288, 284685, 284768, 284768, 284769, 21908768, 284691, 284692, 284693, 18108370, 284628, 284629, 18108374, 284630, 284631, 284634, 284636, 284637, 284638, 18108382, 83373044, 18108385, 264483, 284584, 284585, 284586, 284488, 284587
1965	94317445 (3929, 3930)	Novel Protein sim. GBank gi4107017[db][BAA38284] - (AB001773) PEM-8 [Clona savignyi]	ubiquitin	284488, 284510, 264760, 264768, 264486

1966	94192058 (3931, 3932)	Novel Protein sim. GBank gi 4929707 gb AAD34114.1 AF15187 - (AF151877) CGI-119 protein [Homo sapiens]	Contains protein domain (PF01027) - Uncharacterized protein family	glycoprotein	22278999, 264092, 264259, 29331828, 29331828, 29146488, 264595, 285011, 264448, 18108354, 264288, 264684, 264766, 264685, 264688, 265022, 264691, 264692, 18108370, 18108377, 264555, 18108381, 18108385, 264486, 264587
1967	87396123 (3933, 3934)	Novel Protein sim. GBank gi 2957270 (AF044576) - phospholipase C PLC210 [Caenorhabditis elegans]	Contains protein domain (PF00388) - Phosphatidylinositol-specific phospholipase C, X domain	esterase	29331824, 265010, 265017, 264288, 21906764, 263981, 56528486
1968	88095841 (3935, 3936)	Novel Protein sim. GBank gi 2564953 (AF030001) - unknown [Mus musculus]	Contains protein domain (PF00008) - EGF-like domain	oncogene	35696286, 264805, 264508, 264808, 264807, 264808, 264909, 264511, 264512, 265008, 264910, 265009, 264594, 264757, 264758, 264604, 264605, 264760, 264762, 264682, 264764, 264685, 264768, 264767, 264689, 264691, 264693, 264628, 264629, 35696423, 35695955, 264631, 264632, 264634, 264635, 264638, 264637, 18108380, 264584, 264585, 264586, 264587
1969	84326529 (3937, 3938)	Novel Protein sim. GBank gi 2911274 (U20328) - spidroin 1 [Nephila clavipes]		UNCLASSIFIED	22278995, 22278986, 35696052, 264808, 264908, 18109351, 264482
1970	80596049 (3939, 3940)	Novel Protein sim. GBank gi 4050087 (AF109907) - S164 [Homo sapiens]		UNCLASSIFIED	264808, 264288, 264766, 264638
1971	94843814 (3941, 3942)	Novel Protein sim. GBank gi 134206 sp P09593 SANT_PLAFV - S-ANTIGEN PROTEIN PRECURSOR		collagen	264488, 264489, 22278988, 264259, 60432049, 66714117, 29331828, 60432289, 29331827, 35698052, 264508, 264805, 264508, 264908, 264807, 264908, 264809, 264510, 264511, 264512, 264610, 264591, 264592, 60432229, 60433356, 264595, 264596, 264600, 264604, 264605, 264760, 18108351, 264448, 264764, 264288, 264768, 264769, 21908765, 33657023, 264692, 18108370, 264629, 35698423, 65274791, 35695855, 264632, 264635, 264555, 264638, 264637, 264638, 264639, 18108385, 60432113, 22278000, 264583, 264584, 264585, 264586, 264488
1972	87645444 (3943, 3944)	Novel Protein sim. GBank gi 4519623 dbj BAA75671.1 - (AB017616) homologous to the yeast YGR163 gene [Mus musculus]	Contains protein domain (PF01462) - Leucine rich repeat N-terminal domain		22278999, 264259, 29331822, 56182181, 60432289, 29331827, 52644045, 264909, 265008, 264511, 265008, 52844296, 265018, 265019, 264761, 264689, 21906768, 21908769, 264691, 264693, 33657109, 33657182, 264556, 52644332, 264558, 60432113
1973	86395533 (3945, 3946)			UNCLASSIFIED	29331828, 264692, 35696423, 264631, 264555, 264556, 264557, 264558, 264559
1974	80396629 (3947, 3948)	Novel Protein sim. GBank gi 3309543 (AF036382) - MLL [Fugu rubripes]		UNCLASSIFIED	264682, 264764, 264583

1975	94316479 (3949, 3950)			UNCLASSIFIED	264488, 66714117, 29331826, 29331828, 56182435, 265006, 264757, 55812038, 265010, 265017, 264369, 55811957, 65274791, 35895855, 56182323, 60432113
1976	95358914 (3951, 3952)			UNCLASSIFIED	264259, 35896052, 265018, 265020, 265021, 33657109, 56526486
1977	94852664 (3953, 3954)	Novel Protein sim. GBank gi2499328[sp]Q07782[NASU_RAT - SODIUM/SULFATE COTRANSPORTER (NA(+)/SULFATE COTRANSPORTER)]		homeobox	264808, 264596, 265021, 264566
1978	87447645 (3955, 3956)	Novel Protein sim. GBank gi103421[pir]A33471 - transcription factor NTF1 - fruit fly (Drosophila melanogaster) (fragment)		transcription factor	60170831, 264568
1979	87627709 (3957, 3958)	Novel Protein sim. GBank gi2244815[emb]CAB10238.11 - (Z97336) hypothetical protein (Arabidopsis thaliana)		UNCLASSIFIED	29331828, 29146498, 264905, 264807, 265007, 265009, 265010, 265018, 264886, 18108359, 21908768, 35895917, 265020, 60170815, 264693, 18108388, 18108370, 264631, 264635, 264558, 264558, 18108384, 22279000, 264565
1980	86577059 (3959, 3960)	Novel Protein sim. GBank gi4759280[ref]NP_004842.1[USP1 - Ubiquitin carboxyl-terminal hydrolase, X-linked]		ubiquitin	264489
1981	87606974 (3961, 3962)	Novel Protein sim. GBank gi4066893[gb]AAD200501 - (AF131849) Unknown [Homo sapiens]		UNCLASSIFIED	22278995, 22278997, 22278999, 264259, 29331825, 29331828, 29146498, 29146499, 264107, 264908, 264910, 264595, 21908754, 265010, 265017, 265018, 265019, 264448, 264288, 21908767, 33657023, 27486264, 18108370, 18108374, 18108376, 264630, 264631, 264635, 18108385, 87168518, 22279000, 264482, 264564
1982	80995387 (3963, 3964)	Novel Protein sim. GBank gi5689523[dbj]BAA83045.11 - (AB029016) KIAA1093 protein [Homo sapiens]			65274572, 29331825, 35896052, 33656970, 264909, 265008, 55811386, 264760, 264686, 264691, 27486264
1983	95098688 (3965, 3966)	Novel Protein sim. GBank gi3417297 (AC002310) - Unknown gene product [Homo sapiens]	Contains protein domain (PF000096) - Zinc finger, C2H2 type	transcription factor	22278995, 35896288, 264259, 29331822, 20281099, 29331824, 60432289, 29331827, 264509, 264905, 264906, 264907, 68712502, 264908, 52644045, 264909, 264510, 264512, 264910, 265009, 264591, 264592, 60433356, 60433438, 264758, 265010, 264600, 264603, 264604, 264760, 264762, 264763, 264764, 264766, 264687, 264768, 264769, 21908765, 55811957, 35895917, 264690, 264692, 264693, 264628, 264629, 263978, 18108379, 35896423, 35895855, 20281071, 264632, 264634, 264635, 264555, 264636, 264637, 264638, 264639, 83373044, 18108385, 22279000, 22279002

1984	85760989 (3987, 3968)	Novel Protein sim. GBank gij286893[emb CAA17174.1]- (AL021897) fadD14 [Mycobacterium tuberculosis]		synthase	284688, 21906766, 55811957, 56994073, 265020, 265021, 22278999, 265022, 284259, 29331822, 33857182, 29146499, 284628, 18108370, 284908, 284629, 55811576, 35895855, 265008, 265007, 284591, 21908754, 33857084, 265010, 265017, 265019, 264288
1985	85636897 (3988, 3970)	Novel Protein sim. GBank gij5712131[gb AAD47379.1 AF12049] DEM1 protein [Homo sapiens]		glycoprotein	284760, 264288, 263978, 55811576, 284637, 58182323, 18108385, 264584
1986	80200507 (3971, 3972)	Novel Protein sim. GBank		UNCLASSIFIED	284488, 264629
1987	87011117 (3973, 3974)	Novel Protein sim. GBank gij4868443[gb AAD31319.1 AF14457] Mx- Interacting protein kinase PKM [Mesocricetus auratus]	Contains protein domain (PF00089) - Eukaryotic protein kinase domain		22278999, 29331830, 265007, 265018, 21908768, 33857023, 284692, 284693, 18108377, 284635, 60170394, 22278002
1988	94122108 (3975, 3976)			UNCLASSIFIED	284903, 264906, 284907, 264908, 264909, 284910, 264591, 284593, 284758, 284784, 284686, 284788, 285021, 284692, 284628, 284629, 35895855, 284630, 284635, 284636, 284637, 284638, 284639, 264483
1989	91225225 (3977, 3978)	Novel Protein sim. GBank gij2801701 (AF042379) - spindle pole body protein spc97 homolog GCP2 [Homo sapiens]		tubulin	60432049, 60432289, 52844045, 56182435, 284112, 265007, 33857402, 52644229, 21908785, 21908788, 21908789, 55811957, 33857023, 263987, 33857109, 18108370, 22279000, 22279002
1990	85698888 (3979, 3980)	Novel Protein sim. GBank gij5701727[dbj BAA83074.1]- (AB024729) alpha-1,3-D-mannoside beta-1,4-N- acetylglucosaminyltransferase IV-homologue [Homo sapiens]			284508, 264757, 264764, 18108381
1991	85353114 (3981, 3982)	Novel Protein sim. GBank gij4240287[dbj BAA74922.1]- (AB020708) KIAA0899 protein [Homo sapiens]	Contains protein domain (PF01602) - Adaptin N terminal region	glycoprotein	18108394, 56182575, 22278994, 35898288, 56994075, 22278997, 22278999, 29331822, 29331824, 29331825, 60432289, 29331828, 284508, 284908, 284907, 284908, 58182435, 284510, 265007, 21906754, 33109954, 87168474, 265017, 265018, 265019, 284762, 18108351, 284763, 284883, 284369, 284288, 284685, 284766, 284687, 284769, 21906785, 21908768, 21908769, 55811957, 265020, 60431528, 263974, 18108379, 35895855, 284535, 264557, 264639, 83373044, 18108384, 87168518, 60432113, 22278000, 22279002, 264564, 264488
1992	95317232 (3983, 3984)	Novel Protein sim. GBank gij2246532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma- associated herpesvirus]	Contains protein domain (PF00089) - Zinc finger, C2H2 type	struct	29331827, 284908, 284907, 264909, 265007, 284603, 284766, 284886, 264768, 21908768, 284628, 284635, 284636, 18108385, 56526486, 284568, 264567
1993	80054763 (3985, 3986)	Novel Protein sim. GBank gij2563091 (U80761) - C1G26 alternate open reading frame [Homo sapiens]		UNCLASSIFIED	284592, 35898423

1994	94329114 (3987, 3988)	Novel Protein sim. GBank gi 5630077 gb AAD45622.1 AC006017) similar to ALR; similar to AAC51735 (PID:g2358287) [Homo sapiens]	Contains protein domain (PF00856) - mapolymerase SET domain	264488, 22278997, 22278998, 22278999, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 264905, 284907, 284908, 265008, 60432228, 33657402, 60433356, 284757, 60433438, 284758, 33109954, 285011, 285017, 285018, 285019, 264684, 264369, 264685, 264686, 264768, 21908765, 21908767, 21908768, 21908769, 285020, 264690, 18108382, 284693, 85274620, 18108370, 284635, 284555, 284556, 284557, 56182323, 83373044, 59528488, 22279000, 22279002, 264584 264488, 18108396, 22278994, 56994075, 22278996, 35698286, 22278997, 22278999, 264259, 29147620, 58182181, 29331824, 60432289, 29331826, 29331827, 35696052, 29146499, 284905, 284907, 66712502, 58182435, 265008, 265007, 265008, 265009, 60431735, 60433356, 33657402, 284595, 55812038, 33657084, 55811386, 85658542, 265010, 265011, 265017, 265018, 265019, 264761, 284762, 284448, 264683, 264764, 284288, 284768, 284886, 284788, 284789, 56181562, 264689, 21908765, 21908768, 21908767, 28148627, 21908768, 21908769, 29148629, 29148784, 265020, 265021, 264690, 18108361, 264693, 27488262, 27488264, 27488265, 18108370, 60431528, 18108374, 18108377, 35698423, 55811578, 65274781, 35695855, 284631, 284634, 284635, 284555, 284636, 60431850, 284557, 284558, 284559, 83373044, 20798451, 87188518, 284404, 60432113, 264567 264584
1995	85414353 (3989, 3990)	Novel Protein sim. GBank gi 4827040 ref NP_005110.1 pTRAP - thyroid hormone receptor-associated protein, 150 kDa subunit	UNCLASSIFIED	
1996	80254186 (3991, 3992)	Novel Protein sim. GBank gi 791146 emb CAA60020 - (X88028) extensin-like protein [Vigna unguiculata]	UNCLASSIFIED	
1997	87028423 (3993, 3994)	Novel Protein sim. GBank gi 2642034 (AF034547) - protein phosphatase M130 myosin binding subunit [Ovis aries]	Contains protein domain (PF00023) - phosphatase Ank repeat	264908, 264909, 264592, 264593
1998	85262704 (3995, 3996)	Novel Protein sim. GBank gi 4589634 dbj BAA76639.1 - (AB023212) KIAA0995 protein [Homo sapiens]	UNCLASSIFIED	264113, 264685, 284555, 284567

1899	84324903 (3987, 3988)	Novel Protein sim. GBank gi 5225312 gb AAD0846.1 AF07244 - (AF072441) calchneurin binding protein cabin 1 (Homo sapiens)	Contains protein domain (PF00515) - TPR Domain	UNCLASSIFIED	18108394, 18108397, 35696286, 60424269, 29331827, 29331828, 35696052, 285008, 284512, 55811386, 265010, 265018, 265019, 55811150, 18108351, 264763, 284682, 284389, 284685, 284686, 56181582, 265020, 284691, 33657023, 284693, 33657109, 27486284, 18108370, 18108379, 35695855, 284634, 284635, 284636, 284555, 284557, 58182323, 18108382, 284559, 83373044, 60432113, 22279000, 284563, 284564, 284566
2000	95413705 (3999, 4000)	Novel Protein sim. GBank gi 1723232 sp Q10155 YATA_SCHPO - HYPOTHETICAL 90.6 KD PROTEIN C1D4.10 IN CHROMOSOME 1	UNCLASSIFIED	UNCLASSIFIED	52646365, 52646842, 22278994, 22278995, 22278998, 35696286, 22278997, 22278998, 22278999, 264259, 52645080, 29331822, 29331828, 29331827, 29331828, 35696052, 284108, 28331830, 52644045, 265007, 265008, 60170831, 284592, 284593, 33657402, 60433438, 21906754, 52644286, 285017, 285018, 285019, 264761, 284369, 284288, 52644229, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265020, 285021, 60170815, 52644150, 33657023, 65274620, 52645129, 27486281, 27486282, 27486284, 35695763, 35696423, 35695855, 284631, 52644332, 58182323, 60170394, 83373044, 56526488, 22279002, 284568, 284567
2001	95072534 (4001, 4002)	Novel Protein sim. GBank gi 107560 pir B38637 - Ras inhibitor (clone JC285) - human (fragment)	UNCLASSIFIED	UNCLASSIFIED	284769, 52644229, 65274572, 21906768, 22278998, 35696286, 35695917, 265020, 22278999, 284534, 264490, 264259, 284692, 60432289, 33657109, 35696032, 284508, 284509, 18108370, 60431528, 18108374, 35696423, 65274791, 35695855, 284510, 284511, 284512, 265009, 284634, 284636, 284555, 284556, 284638, 284557, 284558, 284559, 60433436, 83373044, 284759, 18108385, 265011, 284600, 264601, 60432113, 284603, 284604, 284605, 264448, 284288, 284785
2002	80236368 (4003, 4004)	Novel Protein sim. GBank gi 729433 sp P36657 ER60_BOVIN - PROBABLE PROTEIN DISULFIDE ISOMERASE ER-60 PRECURSOR (ERP60) (58 KD MICROSOMAL PROTEIN) (P58)	Contains protein domain (PF00085) - Thioredoxin	Isomerase	284907, 265008, 284910, 284603, 284692, 284629, 18108374, 284558, 284557
2003	80074449 (4005, 4006)	Novel Protein sim. GBank gi 86388 pir A27040 - neurofilament triplet M protein - chicken (fragment)	UNCLASSIFIED	UNCLASSIFIED	284805, 284906, 264908, 264910, 284596, 265017, 18108351, 284692, 284629, 284634, 284565

2004	95317318 (4007, 4008)	Novel Protein sim. GBank gl[4884249]emb[CAB43230.1] - (AL049990) hypothetical protein [Homo sapiens]	Contains protein domain (PF00070) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	dna_rna_bind	52845158, 52846842, 52846365, 56182575, 22278994, 22278995, 56994075, 22278996, 35698286, 22278997, 22278998, 22278999, 60432049, 284259, 52845080, 29331824, 29331828, 29331827, 35696052, 29331828, 33856970, 29331830, 284908, 284592, 60433356, 33657402, 52846317, 21906754, 33857084, 52844298, 87188474, 87188559, 285017, 265018, 265019, 284763, 264683, 284288, 52844229, 21908765, 21908786, 21908787, 21908788, 21908789, 35695917, 265020, 265021, 285022, 52844150, 33657023, 52845129, 33657109, 33657182, 27486261, 27486282, 33857349, 27486285, 35695763, 18108374, 18108376, 18108377, 35698423, 35695855, 284631, 52844332, 284558, 18108385, 56528486, 87188518, 60432113, 284483, 284488, 264908
2005	87400864 (4009, 4010)	Novel Protein sim. GBank gl[387950]emb[CAB87795] - (Z47812) similar to ubiquitin carboxyl-terminal hydrolase; cDNA EST EMBL:D33366 comes from this gene; cDNA EST EMBL:D33965 comes from this gene; cDNA EST EMBL:D33822 comes from this gene; cDNA EST EMBL:D34547 comes from this ge...		ubiquitin	
2006	95351177 (4011, 4012)	Novel Protein sim. GBank gl[410667]emb[CAB22613] - (AL035064) queuine trna-ribosyltransferase [Schizosaccharomyces pombe]	Contains protein domain (PF01702) - Queuine tRNA-ribosyltransferase	UNCLASSIFIED	56182575, 56181686, 22278995, 22278996, 22278997, 22278998, 22278999, 284259, 56182181, 60424268, 68714117, 35696052, 264808, 68712502, 284909, 284510, 60433356, 85658542, 265010, 265018, 265019, 284882, 284448, 284288, 284768, 29148627, 21906769, 29148784, 35695917, 60170815, 284691, 33657023, 65274620, 33657109, 55810764, 55811576, 35695855, 87188518, 60432113, 284563, 284482
2007	94325556 (4013, 4014)	Novel Protein sim. GBank gl[266216]dbj[BAA23712] - (AB007900) HH0452 cDNA clone for KIAA0440 has a 438-bp insertion at position 1711 of the sequence of KIAA0440. [Homo sapiens]		UNCLASSIFIED	264488, 263994, 35696052, 264508, 264905, 264906, 264907, 264908, 264909, 265007, 264910, 264592, 264595, 264758, 265011, 264760, 284762, 284764, 284768, 284685, 264767, 264768, 284769, 55811957, 35695917, 265020, 284691, 264893, 264628, 264629, 65274791, 35695855, 284631, 284632, 264634, 264635, 284637, 284638, 264639, 284566
2008	85084428 (4015, 4016)	Novel Protein sim. GBank gl[1550783]emb[CAB69257] - (Y07960) homeodomain protein [Mus musculus]	Contains protein domain (PF00046) - Homeobox domain	homeobox	264909, 264768, 35695855

2009	85748240 (4017, 4018)	Novel Protein sim. GBank gi 3882305 dbj BAA34512.1 - (AB018335) KIAA0792 protein [Homo sapiens]	UNCLASSIFIED	22278998, 264259, 264910, 264591, 265017, 264681, 264683, 21908768, 264691, 33657182, 33657349, 264631, 87168518, 264404, 22279002, 264563
2010	85422458 (4019, 4020)	Novel Protein sim. GBank gi 5262629 emb CAB45753.1 - (AL080164) hypothetical protein [Homo sapiens]	eph Contains protein domain (PF00057) - Low-density lipoprotein receptor domain class A	52844507, 52844516, 52846355, 52846842, 18108397, 65274572, 22278994, 56994075, 35688286, 22278998, 22278997, 22278999, 264259, 29331822, 52845080, 29331824, 29331825, 29331826, 29331827, 29331828, 264511, 265007, 264512, 265008, 265009, 60432229, 60433356, 21908754, 52846317, 33108954, 52844296, 87188474, 87168559, 265017, 265018, 265019, 264881, 264885, 264687, 52844229, 264689, 21908785, 21908768, 21908767, 21908768, 35685917, 265020, 52844150, 264691, 264692, 33657023, 263987, 52845129, 35693783, 18108376, 35698423, 65274791, 35695855, 264631, 264634, 60431850, 264637, 264638, 52844332, 60170394, 18108365, 87168518, 22279002, 264584, 264585, 264568, 264567, 56182575, 56994075, 22278999, 264259, 29331824, 29331826, 29331827, 29331828, 35698052, 264906, 66712502, 265008, 265007, 265008, 265010, 265011, 265017, 265019, 264681, 264448, 264683, 264369, 264288, 264685, 264768, 264687, 21908785, 21908767, 21908768, 21908788, 265020, 265022, 264691, 33657023, 65274620, 33657109, 264629, 264557, 264559, 83373044, 87188518, 60432113, 22279002
2011	94328149 (4021, 4022)	Novel Protein sim. GBank gi 3347953 (AF076183) - cytosolic sorting protein PACS-1a [Rattus norvegicus]	UNCLASSIFIED	264259, 29331822, 29331824, 29331825, 28148498, 87168559, 265019, 264682, 284288, 264686, 21908764, 265020, 265022, 33657023, 264693, 33657109, 55811576, 264632, 264558, 56182323, 264639, 18108383, 18108384, 18108386, 22279000, 22279002, 264567
2012	87772137 (4023, 4024)	Novel Protein sim. GBank gi 1086678 (U41020) - coded for by C. elegans cDNA yk100g4.5; coded for by C. elegans cDNA yk100g4.3; weakly similar to human SREBP-2 basic-helix-loop-helix-leucine zipper transcription factor [Caenorhabditis elegans]	UNCLASSIFIED	18108398, 264908, 265007, 265010, 265018, 265019, 264689, 21908767, 265020, 264692, 264639, 264558
2013	84843842 (4025, 4026)	Novel Protein sim. GBank gi 4507885 ref NP_003427.1 pZNF1 - zinc finger protein 135 (clone pHZ-17)	dna_ma_bind	264688, 29331826, 264907, 264638, 264555, 264639, 264558
2014	87347940 (4027, 4028)	Novel Protein sim. GBank gi 127720 sp P20938 MYPO_HETFR - MYELIN PO PROTEIN PRECURSOR	UNCLASSIFIED	

2015	88094922 (4029, 4030)	Novel Protein sim. GBank gl 81286 pir S22697 - extensin - Volvox carter (fragment)		UNCLASSIFIED	56182575, 35696286, 264259, 35696052, 264508, 264906, 264907, 264510, 264512, 87188474, 265010, 264681, 264288, 264689, 264628, 35698423, 35695855, 264639, 264563, 264584
2016	85298641 (4031, 4032)	Novel Protein sim. GBank gl 285046 pir S26413 - t-complex protein Top-10 - mouse		strut	264102, 264508, 264110, 265009, 33109954, 21906768, 265021, 33657109, 27486282, 263972, 18108374, 263978, 264555, 264564
2017	79464293 (4033, 4034)	Novel Protein sim. GBank gl 124735 pir P18175 INVO_PIG - INVOLUCRIN		UNCLASSIFIED	264685, 264636
2018	79637067 (4035, 4036)	Novel Protein sim. GBank gl 2143910 pir S68216 - phosphatase-1 glycogen-binding (GL)-chain - rat		phosphatase	264107, 264110, 264112, 265017, 263976
2020	94674476 (4039, 4040)	Novel Protein sim. GBank gl 2078483 (U43200) - antifreeze glycoprotein precursor [Boreogadus saida]		UNCLASSIFIED	264259, 264508, 264591, 265018, 264682, 264288, 264686, 22279002
2021	86718818 (4041, 4042)	Novel Protein sim. GBank gl 585084 pir Q07803 EFQM_RAT - ELONGATION FACTOR G, MITOCHONDRIAL PRECURSOR (MEF-G)			56994075, 264593, 33109954, 21908754, 21908768, 33657023, 33657109, 27486281, 87188518
2022	95295665 (4043, 4044)	Novel Protein sim. GBank gl 4218005 (AC006135) - putative vicilin storage protein (globulin-like) [Arabidopsis thaliana]			264757, 264767, 60170615, 18108385
2023	87722976 (4045, 4046)	Novel Protein sim. GBank gl 5410230 gb AAD2992.1 AF07334 - (AF07334) ubiquitin specific protease 3 [Homo sapiens]	Contains protein domain (PF00442) - Ubiquitin carboxyl-terminal hydrolases family 2	ubiquitin	18108394, 22278999, 264259, 264905, 264906, 264908, 264595, 264762, 264769, 264634, 264636, 87188518, 60432113, 22279000, 264482, 264565
2024	87896443 (4047, 4048)				80433438, 265017, 264686, 264692, 264893, 264636
2025	87856863 (4049, 4050)	Novel Protein sim. GBank gl 4755168 gb AAD29055.1 AC00701 - (AC007018) unknown protein [Arabidopsis thaliana]	Contains protein domain (PF00637) - 7-fold repeat in Clathrin and VPS	UNCLASSIFIED	22278997, 264509, 264906, 264909, 55812038, 265017, 265021, 265022, 60170815, 264556
2026	84122114 (4051, 4052)	Novel Protein sim. GBank gl 165569 emb CAA69032 - (Y07752) perophorin-S [Volvox carter]		UNCLASSIFIED	56994075, 60432049, 264508, 66712502, 264112, 60170831, 87188559, 264288, 264688, 264689, 21906768, 33657109, 18108370, 264638, 18108385, 60432113, 22279000, 22279002, 264564, 264568, 264567
2027	80249001 (4053, 4054)			UNCLASSIFIED	263978, 264634, 264486

2028	94698864 (4055, 4056)	Novel Protein sim. GBank gi1330345 (U58755) - coded for by C. elegans cDNA yk34b1.5; coded for by C. elegans cDNA yk13h10.5; coded for by C. elegans cDNA yk46a8.5; coded for by C. elegans cDNA yk46d5.5; coded for by C. elegans cDNA yk43c2.5; coded for by C. elegans cDNA yk46e8....	Contains protein domain (PF00632) - ubiquitin-HECT-domain (ubiquitin-transferase).	52644507, 52645158, 52646842, 56182575, 56994075, 35696286, 22278997, 22278998, 22278999, 60432049, 264259, 52645080, 29331822, 29331824, 66714117, 29331825, 29331826, 29331827, 29331828, 35696052, 264908, 264907, 264908, 29331830, 52644045, 56182435, 265006, 265009, 60432228, 33657402, 264595, 264757, 55812038, 21906754, 52646317, 52644286, 265010, 265011, 67168559, 265017, 265018, 265019, 264448, 18108354, 264288, 264389, 264768, 52644228, 21906765, 21906768, 21906767, 21906768, 55811857, 35695917, 265020, 265021, 265022, 52644150, 33657023, 65274620, 33657109, 52645129, 18108368, 27486261, 27486282, 27486284, 27486285, 35695763, 264629, 55811576, 35696423, 35695855, 264635, 264636, 52644332, 264558, 83373044, 56528488, 22279000, 22279002, 284583, 265009, 264595, 55855542, 264555, 284556, 264557, 264558, 264559, 83373044
2029	95362032 (4057, 4058)	Novel Protein sim. GBank gi13599940 (AF017368) - (ecodgenital dysplasia protein 2 [Mus musculus])	Contains protein domain (PF00621) - UNCLASSIFIED RhoGEF domain	18108384, 56994075, 22278997, 22278998, 264259, 29331822, 29331824, 29331825, 66714117, 80432289, 29331826, 264108, 66712502, 264628, 265009, 265018, 265019, 264681, 264682, 264684, 264685, 56181562, 264689, 21906769, 265022, 264692, 264693, 264628, 18108370, 18108374, 264634, 264638, 264558, 18108385, 87188518, 22279002, 264565
2030	91213734 (4059, 4060)	Novel Protein sim. GBank gi15630080 (gb AAD45625.1 AC004890) similar to HUB1; similar to BAA24380 (PID:g2789430) [Homo sapiens]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	264591, 55811957, 18108385, 264557, 264558, 18108382, 18108384
2031	80245281 (4061, 4062)			
2032	91232607 (4063, 4064)	Novel Protein sim. GBank gi15688491 (gb BAA83029.1 - (AB028000) KIAA1077 protein [Homo sapiens])	Contains protein domain (PF00884) - hydrolase Sulfatase	65274572, 35696286, 29331824, 264908, 265009, 264593, 265018, 264288, 264688, 264768, 21906766, 21906767, 29146827, 264628, 35698423, 264634, 264558, 18108381, 60170394, 264559, 83373044, 18108385, 264482, 264484
2033	95000809 (4065, 4066)	Novel Protein sim. GBank gi2494828 (gb O64886 CAG7_RAT - ALPHA-N-ACETYLGLACTOSAMINIDE ALPHA-2,6-SIALYLTRANSFERASE (ST6GALNAC1)) (STY)	synthase	56181562, 264628, 264632, 264555, 264556
2034	91232528 (4067, 4068)	Novel Protein sim. GBank gi14828884 (ref NP_005147.1 PROD1 - UNKNOWN)	Contains protein domain (PF00076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain)	22278996, 22278999, 264907, 29331830, 265008, 265018, 264681, 264682, 264684, 21906767, 21906768, 21906769, 33657109, 83373044, 56528488

2035	83553451 (4059, 4070)				264369, 264686, 265022, 56526488, 264567
2036	87115833 (4071, 4072)				29331827, 29331828, 264682, 284369, 29148827, 60432113
2037	94324833 (4073, 4074)	Novel Protein sim. GBank gij2734081 (AF000195) - similar to oxysterol-binding proteins (Caenorhabditis elegans)		UNCLASSIFIED	65274572, 22278995, 22278996, 56994075, 35698286, 22278997, 22278998, 22278999, 264259, 29331824, 60432289, 29331828, 29331828, 35698052, 264907, 29331830, 66712502, 58182435, 265008, 265009, 60170831, 264594, 55812038, 33109854, 21908754, 87168559, 265017, 265018, 265019, 264762, 264369, 264288, 21906765, 21908767, 21906768, 21906769, 55811957, 35695917, 265020, 265021, 265022, 52844150, 33657023, 33657109, 33657182, 35695763, 35695855, 264632, 264634, 264638, 58182323, 83373044, 60432113, 22278900, 22279002, 264563
2038	95422384 (4075, 4076)	Novel Protein sim. GBank gij3880625[emb CAB07858] - (Z83785) predicted using GeneFinder; similar to RNA recognition motif. (aka RRM, RBD, or RNP domain); cDNA .ST EMBL:T01682 comes from this gene; cDNA EST EMBL:M75823 comes from this gene; cDNA EST EMBL:D27559 comes from this ge...	Contains protein domain (PF01412) - Putative GTP-ase activating protein for Arf	UNCLASSIFIED	22278995, 22278996, 56994075, 264259, 29331824, 35698052, 264905, 264906, 52644045, 265007, 265009, 87168559, 265017, 18108351, 264448, 264369, 284768, 264767, 264688, 18108358, 21906765, 21906769, 52844150, 33657023, 264692, 18108362, 33657109, 27486262, 18108370, 18108374, 18108379, 35698423, 65274791, 264632, 264638, 18108383, 83373044, 18108385, 87168518, 22279000, 22279002, 264563, 264564, 284586
2039	85514628 (4077, 4078)	Novel Protein sim. GBank gij2224653[dbj BAA20813] - (AB002354) KIAA0356 [Homo sapiens]	Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger)	UNCLASSIFIED	22278997, 264259, 29331822, 264905, 264906, 264907, 264908, 264909, 264510, 265008, 264910, 264593, 264758, 265011, 265018, 264762, 264288, 264768, 264768, 264769, 21906766, 33657023, 264692, 264693, 33657109, 35698423, 264631, 264632, 264634, 264635, 264636, 264637, 264639, 87168518, 284486
2040	85308417 (4079, 4080)			UNCLASSIFIED	284592
2041	95071736 (4081, 4082)	Novel Protein sim. GBank gij2500625[sp P70700 P2A_MOUSE - DNA-DIRECTED RNA POLYMERASE I 135 KD POLYPEPTIDE (RNA POLYMERASE I SUBUNIT 2) (RPA135)	mapolymerase		264488, 22278998, 35698052, 264905, 284907, 264908, 264910, 265018, 264605, 265019, 18108351, 264768, 264769, 21908768, 265021, 285022, 264692, 33657109, 264628, 264629, 35698423, 35695855, 264637, 264638, 264563, 264584, 264565, 264567

2042	95307447 (4083, 4084)	Novel Protein sim. GBank gij4406590jgb/AAD20040] - (AF131768) Similar to Ena-VASP like protein [Homo sapiens]	Contains protein domain (PF00568) - WH1 domain	UNCLASSIFIED	60424178, 35696286, 264259, 29331826, 35696032, 29331828, 264508, 264509, 284907, 264909, 264510, 264511, 265009, 284910, 264591, 60433358, 264595, 285017, 265019, 264681, 264764, 264369, 264765, 264884, 264288, 264766, 264686, 52644229, 264769, 21906765, 35695917, 264535, 52644150, 264691, 264692, 18108385, 27486261, 27486262, 27486265, 18108374, 35696423, 65274791, 35695855, 264555, 264558, 60170394, 18108385, 264404, 22279000, 22279002, 264482, 264583, 264584, 264586
2043	94328078 (4085, 4086)	Novel Protein sim. GBank gij5052554jgb/AAD38607.1jAF14563 - (AF145632) BcDNA, GH06032 [Drosophila melanogaster]	Contains protein domain (PF00122) - E1-E2 ATPase	transport	264488, 52644507, 52646365, 56994075, 22278997, 22278999, 20281171, 264259, 29331822, 29331824, 66714117, 29331826, 29331828, 33656970, 29146498, 264509, 264908, 52644045, 56182435, 265006, 33657402, 21906754, 52644298, 87168559, 265017, 265018, 265019, 264881, 264288, 264766, 264685, 264686, 21906768, 21906767, 21906788, 21906789, 265020, 265021, 60170815, 264691, 33657023, 264693, 65274620, 33657109, 33657182, 27486261, 27486262, 33657349, 35695763, 18108374, 55611576, 35695855, 18108380, 18108381, 60170394, 56182323, 264558, 83373044, 18108385, 56526486, 87168518, 60432113, 22279000, 264567
2044	67106927 (4087, 4088)	Novel Protein sim. GBank gij2246532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus]		struct	264093, 29331827, 264905, 66712502, 264592, 264689, 21906765, 21906769, 265020, 264692, 264482, 264566
2045	78635532 (4089, 4090)	Novel Protein sim. GBank gij4406699jgb/AAD20062] - (AF131852) Unknown [Homo sapiens]			264692
2046	87320849 (4091, 4092)	Novel Protein sim. GBank gij4406699jgb/AAD20062] - (AF131852) Unknown [Homo sapiens]			264258, 264808, 264683, 22278002
2047	84576801 (4093, 4094)	Novel Protein sim. GBank gij4101720 (AF008466) - lymphocyte specific formin related protein [Mus musculus]			22278999, 29147620, 29331824, 29146498, 264508, 265007, 265008, 265019, 264605, 264681, 29148627, 29148629, 265021, 33657023, 18108385, 33657109, 33657182, 18108377, 264558, 264638, 264559, 18108388
2048	84608378 (4095, 4096)	Novel Protein sim. GBank gij4589656jgb/BAA76850.1] - (AB023223) KIAA1008 protein [Homo sapiens]		UNCLASSIFIED	264909
2049	88094690 (4097, 4098)	Novel Protein sim. GBank gij4589656jgb/BAA76850.1] - (AB023223) KIAA1008 protein [Homo sapiens]		UNCLASSIFIED	264488, 264259, 29331824, 29331828, 35696052, 264908, 264907, 264808, 264909, 284910, 264603, 264763, 21906767, 21906768, 264629, 264634, 264637, 22279002, 264584, 264585, 264568, 264567

2050	79633835 (4099, 4100)			UNCLASSIFIED	264693	264488, 264259, 264509, 264906, 264907, 264769, 18108374, 35696423, 264563, 264566, 264488
2051	87760168 (4101, 4102)			UNCLASSIFIED		
2052	88096393 (4103, 4104)	Novel Protein sim. GBank glj4529889[gbjAAD21812.1] - (AF134726) G9A [Homo sapiens]	Contains protein domain (PF00856) - SET domain	kinase		264488, 263994, 35696052, 264508, 264905, 264509, 264906, 264907, 264908, 264909, 264113, 264511, 265009, 264910, 60170831, 264592, 264758, 265010, 265011, 264605, 264760, 264682, 264764, 264369, 264768, 264688, 264768, 264769, 52644229, 264689, 35695917, 33657023, 33657109, 264628, 18108374, 35698423, 55811576, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264556, 264638, 264639, 18108385, 56526486, 60432113, 264563, 264564, 264566, 264488, 264567, 264488, 263994, 35696052, 264508, 264905, 264509, 264908, 264907, 264908, 264909, 264113, 264511, 265009, 264910, 60170831, 264592, 264758, 265010, 265011, 264605, 264760, 264682, 264764, 264369, 264768, 264688, 264768, 264769, 52644229, 264689, 35695917, 33657023, 33657109, 264628, 18108374, 35698423, 55811576, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264556, 264638, 264639, 18108385, 56526486, 60432113, 264563, 264564, 264566, 264488, 264567
2053	87763078 (4105, 4106)	Novel Protein sim. GBank glj2995449[embjCAA75113] - (Y14848) midline 1 protein [Mus musculus]		UNCLASSIFIED		22278996, 22278997, 264259, 29331822, 264102, 264508, 35695917, 263972, 264482
2054	95358937 (4107, 4108)	Novel Protein sim. GBank glj3876326[embjCAB02090] - (Z79754) similar to C2 domain [Caenorhabditis elegans]	Contains protein domain (PF00168) - C2 domain			60424178, 264084, 264259, 29331825, 60424269, 264908, 60432229, 60433356, 87168559, 265019, 264760, 264288, 264686, 21906769, 33657023, 264693, 55810764, 55811576, 264635, 56182323, 60432113
2055	88259448 (4109, 4110)	Novel Protein sim. GBank glj5353746[gbjAAD2226.1]AF15913 - (AF159133) SIR2-like protein [Oryza sativa subsp. indica]		UNCLASSIFIED		264488, 29331826, 60432289, 29331828, 60433356, 265019, 264683, 264684, 265021, 33657109, 18108374, 264637, 18108385, 87168518, 60432113, 22279000, 264564

2056	88177398 (4111, 4112)	Novel Protein sim. GBank gll4826960[rel]NP_005042.1pQARS - glutamine-IRNA synthetase	Contains protein domain (PF00749) - IRNA synthetases class I (E and Q)	- synthase	284488, 52645156, 56182575, 22278994, 35696286, 56994075, 22278998, 22278998, 22278999, 60432049, 264259, 29331824, 60432289, 29331827, 29331828, 33656970, 264104, 264908, 264808, 265008, 265008, 60170831, 264591, 60432228, 60433438, 18108348, 21906754, 33657094, 52844286, 87168474, 285010, 87168559, 265017, 285018, 284760, 18108351, 264681, 264682, 264448, 264683, 264369, 264288, 264685, 264687, 264688, 264689, 21906785, 21906766, 21906787, 21906769, 55811957, 35695917, 265022, 33657023, 18108382, 33657109, 18108368, 33657182, 27486261, 27486264, 27486285, 33657349, 264628, 18108370, 264628, 18108374, 18108377, 18108379, 35698423, 55811576, 20281152, 264636, 264952, 18108385, 18108388, 87168518, 284482, 284585, 284588, 284587, 52846942, 52846365, 56182575, 35696286, 22278998, 22278997, 22278998, 264093, 52845080, 35698052, 29331828, 33656970, 265009, 52846317, 55811386, 52844298, 52844229, 21906789, 35695917, 265021, 60170815, 52844150, 33657109, 33657182, 27486261, 27486282, 35695763, 35698423, 35695855, 52844332 265007, 265008, 264591
2057	87877905 (4113, 4114)	Novel Protein sim. GBank gll728650[sp]P08640[AMYH_YEAST - GLUCOAMYLASE S1/S2 PRECURSOR (GLUCAN 1,4-ALPHA- GLUCOSIDASE) (1,4-ALPHA-D-GLUCAN GLUCOHYDROLASE)		UNCLASSIFIED	
2058	88276898 (4115, 4116)	Novel Protein sim. GBank		UNCLASSIFIED	
2059	78866684 (4117, 4118)	gll119714[sp]P13983[EXTN_TOBAC - EXTENSIN PRECURSOR (CELL WALL HYDROXYPROLINE-RICH GLYCOPROTEIN)]		UNCLASSIFIED	
2060	83050800 (4119, 4120)	Novel Protein sim. GBank gll2811122 (U87318) - NaDC-2 [Xenopus laevis]		UNCLASSIFIED	56182575, 29331824, 29331826, 264910, 55811957, 18108370, 55811576

2061	95362204 (4121, 4122)	Novel Protein sim. GBank gij2496947jsp[Q09288]YQO9_CAEEL - HYPOTHETICAL 141.2 KD PROTEIN EEED8.8 IN CHROMOSOME II	Contains protein domain (PF00069) - kinase Eukaryotic protein kinase domain	22278997, 22278999, 264259, 29331822, 29331824, 29331828, 29331827, 35696052, 29331828, 264906, 66712502, 29331830, 264908, 264909, 264112, 264511, 265007, 265009, 264910, 264591, 33657402, 21906754, 85658542, 265017, 265019, 264448, 264683, 264288, 264684, 264369, 264688, 264687, 18108358, 264689, 21906785, 21906787, 21906788, 21906789, 265020, 265022, 264691, 33657023, 33657109, 20281149, 18108379, 35695855, 264634, 264556, 264557, 264558, 18108382, 264559, 83373044, 18108384, 56528486, 60432113
2062	87028440 (4123, 4124)	Novel Protein sim. GBank gij4502091jre[INP_001139, 1pANK2 - ankyrin 2, neuronal	Contains protein domain (PF00023) - struct Ank repeat	264905, 264628, 264907, 264629, 264908, 264909, 18108374, 263978, 35695855, 264512, 264635, 60431850, 264636, 264760, 264563, 18108351, 264762, 264565, 264764, 264487, 264768
2063	87601272 (4125, 4126)	Novel Protein sim. GBank gij4589562jdbj[BAA76803.1] - (AB023176) KIAA0959 protein [Homo sapiens]	Contains protein domain (PF00617) - oncogene RasGEF domain	22278994, 22278999, 264259, 29331827, 264906, 264909, 52644045, 264686, 21906787, 55811957, 264692, 18108365, 263972, 55811578, 18108384, 22278002, 264482, 264563, 264564, 264484
2064	95317253 (4127, 4128)	Novel Protein sim. GBank gij1764515jdbj[BAA13413.1] - (D87515) aminopeptidase-B [Rattus norvegicus]	hydrolase	264488, 52646365, 56994075, 35696286, 22278997, 22278998, 264259, 29331828, 60432289, 29331827, 29331828, 35696052, 264509, 265007, 265008, 60432229, 60433438, 21906754, 265010, 265011, 87168559, 265017, 265018, 264781, 18108351, 264682, 264369, 264288, 52644229, 21906765, 21906787, 21908768, 35695917, 33657109, 18108368, 18108374, 35698423, 35695855, 52644332, 264559, 60432113, 22279000, 22279002, 264566, 264486
2065	95092238 (4129, 4130)	Novel Protein sim. GBank gij2507144jsp[Q04205]TENS_CHICK - TENSIN	kinase	264569, 18108394, 56182181, 60432289, 29331828, 264905, 264906, 264908, 60431735, 60433356, 55811388, 85658542, 265018, 55811150, 264681, 264788, 264692, 60431528, 263974, 55810764, 35695855, 264631, 264634, 264635, 60431850, 264557, 83373044, 18108388, 22279000, 22279002, 56182575, 264259, 264906, 264784, 264288, 56182323, 264567
2066	85783402 (4131, 4132)	Novel Protein sim. GBank gij160171 (M58295) - circumsporozoite protein [Plasmodium yoelii]	(Contains protein domain (PF00096) - UNCLASSIFIED Zinc finger, C2H2 type	

2067	95303892 (4133, 4134)				35686268, 22278997, 22278998, 60432049, 264259, 60432289, 60433438, 264682, 264448, 264369, 264288, 18108355, 21908765, 21908768, 265022, 33857109, 35686423, 35685855, 264558, 264404, 264553, 264486
2068	84344754 (4135, 4136)			UNCLASSIFIED	264687
2069	94319177 (4137, 4138)	Novel Protein sim. GBank gij3152862 (AF084604) - KE03 protein [Homo sapiens]	Contains protein domain (PF00023) - Ank repeat		60424179, 56182575, 22278995, 22278998, 56994075, 264259, 29331822, 29331824, 29331825, 35686052, 29331828, 33858970, 264509, 264905, 56182435, 265009, 60433358, 87108559, 265017, 265018, 264604, 265019, 264448, 264784, 264766, 21908765, 21908767, 21908768, 21908769, 265020, 265021, 33857023, 33857109, 263976, 264555, 264557, 56182323, 83373044, 87168518, 60432113, 22279000, 22279002
2070	85791380 (4139, 4140)	Novel Protein sim. GBank gij5712131gbjAAD47379 - jAF12049 - (AF120499) DEM1 protein [Homo sapiens]		UNCLASSIFIED	35685917, 264905, 264628, 264908, 264638
2071	86946116 (4141, 4142)	Novel Protein sim. GBank gij3551531dbjBAA330161 - (AB017437) avena [Gallus gallus]	Contains protein domain (PF00588) - WH1 domain		18108398, 265006, 265007, 265008, 265009, 264594, 265010, 265011, 18108351, 18108354, 18108364, 18108365, 18108368, 264634, 18108381, 18108385, 18108388, 18108391
2072	91718428 (4143, 4144)		Contains protein domain (PF00184) - Neurohypophysial hormones, C-terminal Domain		22278995, 35686288, 22278997, 22278998, 22278999, 264490, 60432049, 264259, 29331822, 29331824, 29331828, 35686052, 265008, 33857402, 21908754, 265011, 265019, 18108351, 264682, 264369, 21908765, 21908768, 21908767, 21908768, 21908769, 35685917, 265020, 265021, 264690, 264692, 35686423, 264555, 264556, 264558, 22279000
2073	27925664 (4145, 4146)	Novel Protein sim. GBank gij1504026dbjBAA13212 - (D66978) similar to C.elegans protein (Z37093) [Homo sapiens]		UNCLASSIFIED	264558
2074	84324787 (4147, 4148)	Novel Protein sim. GBank gij4240317dbjBAA74937.1 - (AB020721) KIAA0914 protein [Homo sapiens]			29331822, 264909, 264511, 265009, 264594, 264595, 265010, 265011, 265017, 265018, 265019, 264448, 264883, 265020, 265021, 18108370, 264632, 83373044, 264587

2075	94314886 (4149, 4150)	Novel Protein sim. GBank gij5138930jgb AAD40382.1 - (AF093880) transcription factor IIB [Homo sapiens]			18108394, 22278994, 22278996, 35696286, 22278998, 22278999, 264259, 29331822, 29331825, 29331827, 35696052, 29331828, 284905, 284907, 284908, 284510, 285007, 284910, 285009, 33857084, 284760, 284448, 284288, 284786, 284767, 284689, 21908788, 21808767, 21906769, 265021, 285022, 60170615, 33657023, 27486262, 27486265, 35696423, 35695855, 284631, 284634, 284639, 87168518, 22279002, 284563, 284486, 18108391
2076	87594118 (4151, 4152)				284259, 29331828, 284508, 284908, 284510, 285007, 285011, 284288, 284637, 18108385, 284592
2077	11389877 (4153, 4154)		UNCLASSIFIED		284592
2078	8753384 (4155, 4156)	Novel Protein sim. GBank gij4220590jdb BAA74579 - (D87908) nuclear protein np85 [Mus musculus]	Contains protein domain (PF00828) - PHD-finger		29331825, 285017, 285018, 284288, 285020, 285021, 284634, 56528486
2079	88095916 (4157, 4158)	Novel Protein sim. GBank gij4240255jdb BAA74906.1 - (AB020690) KIAA0883 protein [Homo sapiens]	Contains protein domain (PF00088) - Zinc finger, CCHC class	UNCLASSIFIED	284788, 22278997, 285021, 284690, 284259, 284692, 29331822, 284693, 29331824, 29331828, 284508, 284509, 284908, 284907, 284628, 20281089, 284909, 285007, 285009, 284632, 284638, 284591, 284592, 284639, 284758, 284759, 33108954, 284604, 285018, 285019, 22279002, 284563, 284584, 284448, 284684, 284587, 284685
2080	84136689 (4159, 4160)	Novel Protein sim. GBank gij2408021 emb CAB18219.1 - (Z99162) putative vacuolar protein [Schizosaccharomyces pombe]	Contains protein domain (PF01363) - FYVE zinc finger	UNCLASSIFIED	58182575, 264092, 29331824, 29331826, 29331830, 285017, 285018, 285020, 83373044
2081	94847186 (4161, 4162)	Novel Protein sim. GBank gij5524734jgb AAD44360.1 AF16635 - (AF166350) ST7 protein [Homo sapiens]	Contains protein domain (PF00431) - CUB domain	eph	22278996, 22278997, 22278999, 60432048, 29331822, 29331824, 33657402, 85658542, 285011, 285018, 285019, 21908787, 21908768, 284693, 18108385, 22279000, 22279002
2082	87628629 (4163, 4164)	Novel Protein sim. GBank gij3880558 emb CAA94234 - (Z70271) predicted using Genalinder; similar to collagen; cDNA EST yk308e7.3 comes from this gene; cDNA EST yk308e7.5 comes from this gene; cDNA EST yk385a8.3 comes from this gene; cDNA EST yk385a8.5 comes from this gene [Caeno...]		collagen	284907, 265019
2083	84141000 (4165, 4166)	Novel Protein sim. GBank gij2352427 (AF004161) - peroxisomal Ca-dependent solute carrier [Cryptotagus cuniculus]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	transport	58182575, 22278996, 22278997, 284259, 29331822, 29331823, 284509, 284112, 285009, 284593, 21908754, 265018, 265019, 284448, 284288, 284885, 21908769, 285020, 265022, 284691, 18108370, 65274791, 284631, 284555, 284563
2084	95199298 (4167, 4168)	Novel Protein sim. GBank gij728836jpb P39193 ALU6_HUMAN - III ALU SUBFAMILY SP WARNING ENTRY IIII		UNCLASSIFIED	285018, 284763, 284883, 284691

2085	94889476 (4169, 4170)	Novel Protein sim. GBank gi 165569 emb CAA69032 - (Y07752) pherophorin-S [Volvox carter]	UNCLASSIFIED	56182575, 60432289, 264908, 56182435, 87168474, 264763, 264369, 264688, 264693, 18108370, 56182323
2086	91234404 (4171, 4172)	Novel Protein sim. GBank gi 3875032 emb CAA88936 - (Z49125) similarity to Trichostrogylus colubriformis 11 kd secretory protein (Swiss Prot accession number P21837); cDNA EST EMBL:D33349 comes from this gene; cDNA EST EMBL:D37644 comes from this gene; cDNA EST EMBL:D38149 come...	UNCLASSIFIED	35696286, 264259, 35696052, 264906, 264907, 264908, 264909, 264910, 264759, 264604, 264782, 264768, 264769, 35695917, 263978, 35698423, 35695855, 264632, 264634, 264637, 264638, 264639, 56182323, 18108385, 264482, 264486
2087	21436337 (4173, 4174)		UNCLASSIFIED	264489
2088	84111527 (4175, 4176)	Novel Protein sim. GBank gi 3880930 emb CAA16334.1 - (AL021481) similar to Phosphoglucosylase and phosphomannomutase phosphoserine; cDNA EST EMBL:D36168 comes from this gene; cDNA EST EMBL:D70687 comes from this gene; cDNA EST yk373h9.5 comes from this gene; cDNA EST EMBL:T0080...		264488, 22278994, 35696286, 22278996, 29331827, 35696052, 33857402, 21908754, 33109954, 87168474, 265017, 265018, 265019, 264448, 264683, 264369, 264685, 264687, 264689, 21908765, 21908768, 21908767, 21908768, 21908769, 265020, 265021, 265022, 264692, 33857023, 33857109, 33857182, 27486281, 27486262, 33857349, 27486265, 35698423, 35695855, 83373044, 87168518, 22279000, 264567
2089	95422801 (4177, 4178)	Novel Protein sim. GBank gi 475818 ref NP_004623.1 pDAP3 - Death associated protein 3	cadherin	18108392, 264488, 52844507, 18108394, 18108397, 52848842, 18108398, 56182575, 22278994, 22278995, 35696286, 22278996, 56994075, 22278997, 22278998, 22278999, 264091, 264092, 264093, 264094, 60432049, 264259, 28331822, 20281099, 29331824, 29331825, 29331826, 29331827, 29331828, 35696052, 33856970, 29146498, 29146499, 264102, 264108, 264107, 264109, 264508, 264905, 264509, 264906, 264907, 264908, 66712502, 264828, 52644045, 264909, 56182435, 264110, 264112, 264510, 264511, 265008, 264512, 265007, 265008, 264910, 265009, 60170831, 264592, 264593, 60433356, 33657402, 60433438, 264595, 55812038, 264758, 21908754, 33657084, 55811386, 52844296, 265010, 265011, 87188559, 265017, 265018, 265019, 264760, 264761, 55811150, 264762, 18108351, 264682, 264448, 264763, 264764, 264683, 264369, 18108354, 264258, 264685, 264766, 264688, 264687, 264768, 52844229, 264688, 18108358, 56181562, 264769, 18108359, 264689, 21908765, 21908766, 21908767, 21908768, 29148627, 21908769, 55811957, 29148628, 29148784, 35695917, 265020, 265021, 265022, 60170815, 264690

2080	88222470 (4178, 4180)				22278995, 22278998, 22278999, 264259, 29331828, 35696052, 284910, 33857402, 60433438, 33108954, 87168474, 87168559, 265018, 265019, 264681, 264684, 264688, 264687, 264688, 264689, 21908785, 21908788, 21908787, 21908789, 35695917, 285022, 60170615, 33857023, 35698423, 35695855, 264952, 18108387, 22279000, 263994, 284905, 284908, 264511, 264512, 285008, 264910, 55811386, 264288, 264768, 56181562, 21908785, 21908788, 21908789, 265022, 264628, 284583, 264567
2081	95309161 (4181, 4182)	Novel Protein sim. GBank gl 4380997 gb AAD24571.1 AF12108 - (AF121081) cAMP inducible 2 protein [Mus musculus]	UNCLASSIFIED		
2082	88223805 (4183, 4184)		homeobox		22278997, 22278999, 68712502, 87168559, 284683, 265021, 284486
2093	87406073 (4185, 4188)	Novel Protein sim. GBank gl 2352427 (AF004161) - peroxisomal Ca-dependent solute carrier [Oryctolagus cuniculus]	transport	Contains protein domain (PF00153) - Mitochondrial carrier proteins	284903, 264906, 284907, 284908, 264510, 285008, 265007, 285009, 284910, 284598, 21908754, 87168474, 285011, 284603, 265018, 265019, 284760, 264768, 284788, 284789, 21908787, 21908788, 21908789, 285021, 264690, 33857023, 264693, 284628, 284834, 264638, 284637, 264557, 58182323, 284584
2094	91230928 (4187, 4188)	Novel Protein sim. GBank gl 4929551 gb AAD34036.1 AF15179 - (AF151799) CGI-40 protein [Homo sapiens]	MHC		35696286, 265017, 265018, 265019, 18108388
2095	95351526 (4189, 4190)	Novel Protein sim. GBank gl 1363238 pir JAS7284 - spermatid perinuclear RNA-binding protein Spnr - mouse	dna_rna_bind	Contains protein domain (PF00035) - Double-stranded RNA binding motif	35696286, 52544045, 265006, 265007, 285008, 87168559, 18108351, 21908789, 28148784, 265020, 33857023, 27488282, 18108374, 18108388
2096	94118760 (4191, 4192)	Novel Protein sim. GBank gl 3834423 (AF070689) - cytoplasmic dynein intermediate chain isoform DIC1a [Drosophila melanogaster]	ATPase-associated	Contains protein domain (PF00400) - WD domain, G-beta repeat	284488, 284489, 65274572, 58182575, 22278996, 22278997, 22278999, 264259, 60432289, 29331826, 35696052, 284107, 264508, 264509, 284905, 284906, 284907, 284908, 52844045, 284909, 284510, 284511, 284512, 265008, 264910, 265009, 264592, 80433356, 80433438, 284758, 284598, 55812038, 21908754, 264601, 264602, 264605, 284782, 264681, 18108351, 284784, 284683, 264288, 264687, 264768, 264789, 284689, 21908785, 21908788, 21908787, 35695917, 265020, 265022, 52644150, 264691, 284692, 33857023, 284693, 27488281, 35695783, 264628, 264629, 35696423, 35695855, 264631, 264632, 284634, 264635, 284555, 264637, 263981, 284638, 284639, 284583, 264483, 284585, 264566, 284486, 284567

2087	85322772 (4183, 4194)	Novel Protein sim. GBank gi 5174501 ref NP_008031.1 pLYF1 - zinc finger protein, subfamily 1A, 1 (Ikaro)	Contains protein domain (PF00098) - Zinc finger, C2H2 type	transcriptfactor	65274572, 264511, 265010, 264600, 265017, 264448, 264288, 265021, 60170815, 264692, 33657108, 18108370, 264636, 264483 56994075, 264259, 264288, 265020, 264563
2088	87780340 (4195, 4196)	Novel Protein sim. GBank gi 4759208 ref NP_004081.1 pDUSP - dual specificity phosphatase 3 (vaccinia virus phosphatase VH1-related)	Contains protein domain (PF00782) - Dual specificity phosphatase, catalytic domain	phosphatase	
2089	95412927 (4197, 4198)	Novel Protein sim. GBank gi 2695659 (AF026954) - pyruvate dehydrogenase phosphatase regulatory subunit precursor, PDP1 (Bos taurus)		phosphatase	65274572, 264805, 65274444, 264691, 264636, 264555
2100	85332858 (4199, 4200)	Novel Protein sim. GBank gi 3681189 emb CAB16514 - (Z69281) similar to ADP-ribosylation factor; cDNA EST EMBL: C08179 comes from this gene; cDNA EST EMBL: C08337 comes from this gene; cDNA EST EMBL: C09829 comes from this gene; cDNA EST yk291b4.5 comes from this gene; cDNA EST yk4...	Contains protein domain (PF00025) - ADP-ribosylation factor family	nucd_recp	56182575, 22278895, 22278996, 22278997, 22278998, 60432048, 264259, 29331822, 29331824, 29331825, 29331827, 29331828, 29146498, 264509, 265008, 265009, 264910, 264591, 60432229, 60433358, 33657402, 264758, 21908754, 85658542, 87168474, 265017, 265018, 265019, 264681, 18108351, 264762, 264448, 264369, 264288, 18108355, 264686, 21908765, 21908767, 21908768, 21908769, 265020, 265021, 33657023, 18108374, 35896423, 264558, 83373044, 87188518, 60432113, 22278000, 22278002 264091, 29331824, 264105, 265007, 265010, 18108380
2101	87782604 (4201, 4202)	Novel Protein sim. GBank gi 458468 db BAA76761.1 - (AB012808) mBOCT [Mus musculus]		UNCLASSIFIED	
2102	87770461 (4203, 4204)	Novel Protein sim. GBank gi 3874149 emb CAA87423.1 - (Z73103) predicted using GeneFinder [Caenorhabditis elegans]		UNCLASSIFIED	264488, 264489, 35696286, 264259, 35696052, 264508, 264905, 264907, 264908, 264909, 264511, 264512, 264591, 264593, 60433358, 264758, 264601, 264605, 264780, 18108351, 264448, 264764, 264288, 264787, 264788, 21908768, 35695917, 18108374, 264634, 264555, 264559, 264563, 264482, 264488
2103	95413576 (4205, 4206)	Novel Protein sim. GBank gi 4240159 db BAA74658.1 - (AB020642) KIAA0835 protein [Homo sapiens]	Contains protein domain (PF01530) - Zinc finger, C2HC type	transcriptfactor	65274572, 56994075, 22278999, 264259, 29331824, 29331825, 35696052, 29331828, 68712502, 265009, 60170831, 264595, 33109954, 85658542, 87168459, 265017, 265019, 264448, 21908765, 21908768, 265022, 33657023, 27486262, 33657349, 35695763, 60431528, 18108374, 55811576, 56182323, 18108387, 87168518, 60432113, 264584
2104	85776161 (4207, 4208)			UNCLASSIFIED	264592, 264604, 22278000

2105	94848080 (4209, 4210)	Novel Protein sim. GBank gij1707032 (U80445) - coded for by C. elegans cDNA yk13g5.3; coded for by C. elegans cDNA yk21g8.3; coded for by C. elegans cDNA CEMSE18F; coded for by C. elegans cDNA yk126b1.3; coded for by C. elegans cDNA yk65h8.3; coded for by C. elegans cDNA yk65h8....		UNCLASSIFIED	284488, 56182575, 22278994, 56994075, 22278998, 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 29331826, 60432289, 29331827, 29331828, 284508, 284905, 284509, 284907, 29331830, 52844045, 284510, 284511, 285007, 284512, 285009, 60170831, 60432229, 33857402, 60433356, 284595, 60433438, 284758, 33657084, 87188474, 285010, 87188559, 285017, 285018, 285019, 284782, 18108351, 284684, 18108354, 284288, 284686, 52844229, 18108359, 21908785, 21908788, 21908787, 21908788, 21908789, 35695917, 285020, 285022, 60170815, 52844150, 284692, 33657023, 284693, 33657109, 60431528, 18108374, 65274791, 35695855, 284635, 60170394, 284639, 284558, 18108385, 18108387, 58526486, 87188518, 60432113, 284584, 284586, 284567
2106	83365475 (4211, 4212)				285008, 285019
2107	79822662 (4213, 4214)	Novel Protein sim. GBank gij3881524[emb]CAA93883] - (Z70038) ZK1087.4 [Caenorhabditis elegans]			284908, 284639
2108	94233976 (4215, 4216)	Novel Protein sim. GBank gij3176689 (AC003671) - Contains similarity to ubiquitin carboxyl-terminal hydrolase 14 gb[Z35927 from S. cerevisiae, [Arabidopsis thaliana]		UNCLASSIFIED	284905, 284908, 284907, 284908, 284909, 284758, 285011, 284800, 284801, 284784, 284766, 284767, 284768, 284769, 284693, 284829, 35695855, 284632, 284634, 284635, 284638, 284839, 83373044, 284486
2109	80478719 (4217, 4218)			UNCLASSIFIED	18108348, 284769, 18108370, 18108374, 284555, 284556, 284557, 284558
2110	87729075 (4219, 4220)	Novel Protein sim. GBank gij481043[pir]S37671 - bal2 protein - human		UNCLASSIFIED	284564
2111	87818419 (4221, 4222)			UNCLASSIFIED	284768, 35695917, 284630, 284587, 284486
2112	87293783 (4223, 4224)	Novel Protein sim. GBank gij2143639[pir]I56542 - calmodulin-binding protein - rat	Contains protein domain (PF00069) - struct		284508, 284806, 284591, 284682, 22279002
2113	78941388 (4225, 4226)	Novel Protein sim. GBank gij4426629[gb]AAD20459] - (AF100980) protocadherin [Rattus norvegicus]	Eukaryotic protein kinase domain (PF00028) - cadherin		285008
2114	87889342 (4227, 4228)	Novel Protein sim. GBank gij3327184[db]BAA31660] - (ABD14585) KIAA0885 protein [Homo sapiens]	Cadherin domain		66714117, 29331826, 29331827, 60433438, 55812038, 285017, 285019, 284689, 21908789, 55811957, 285020, 285021, 33657109, 60170394, 284558
2115	90993785 (4228, 4230)	Novel Protein sim. GBank gij4757890[ref]NP_004328.1pc8OR - chromosome 8 open reading frame 1		UNCLASSIFIED	85274572, 284689, 284691, 284692, 60432113

2116	88259387 (4231, 4232)	Novel Protein sim. GBank gij2248532 (U83872) - ORF 73, contains large complex repeat CR 73 (Kaposi's sarcoma-associated herpesvirus)		struct	35686286, 22278999, 56182181, 29331824, 29331825, 29331827, 35696052, 284907, 56182435, 265008, 264591, 55812038, 55811386, 87168558, 264288, 264368, 21906789, 29148628, 33857023, 35695763, 55811576, 35696423, 18108385, 29331824, 264511, 265009, 33109954, 265017, 265018, 264288, 264689, 265020, 264892, 56528486, 284482
2117	87788904 (4233, 4234)	Novel Protein sim. GBank gij2330021 (AF019250) - kinesin-related protein; KRP; Costal2 [Drosophila melanogaster]		struct	264259, 264905, 264907, 264908, 264510, 264511, 265009, 264910, 265010, 264802, 264288, 264768, 264693, 263987, 263972, 264638, 264559
2118	87078894 (4235, 4236)	Novel Protein sim. GBank gij1070307 [prij]B56573 - nuclear pore complex glycoprotein p62 - African clawed frog		glycoprotein	264693, 18108385
2119	86999317 (4237, 4238)	Novel Protein sim. GBank gij4321407 [gbj]AAD15748] - (AF047890) ATP-binding cassette protein M-ABC1 [Homo sapiens]		UNCLASSIFIED	264693, 18108385
2120	87788395 (4239, 4240)	Novel Protein sim. GBank gij4885527 [re]NP_005480.1pNSP3 - novel SH2-containing protein 3	Contains protein domain (PF00017) - Src homology domain 2	eph	264091, 264259, 29331826, 29331828, 265017, 264604, 264288, 264685, 265020, 264691, 18108370, 55810764, 264555, 264638, 60432113
2121	80021375 (4241, 4242)	Novel Protein sim. GBank gij4757728 [re]NP_004886.1pAGTA - angiotensin vasopressin receptor AII/AVP-like		UNCLASSIFIED	264601, 264766, 263978
2122	91230931 (4243, 4244)	Novel Protein sim. GBank gij4829551 [gbj]AAD34038.1AF15179 - (AF151799) CGI-40 protein [Homo sapiens]			18108394, 56182575, 22278997, 29331822, 29331824, 29331825, 29331826, 29331828, 284907, 56182435, 265007, 264910, 265010, 265018, 264688, 265020, 55811576, 284555, 264637, 18108382, 83373044, 18108383, 18108384, 56528486, 264565, 284567, 18108386, 264757, 265011, 18108351, 264691, 264634, 18108385
2123	86787898 (4245, 4246)	Novel Protein sim. GBank gij2224551 [dbj]BAA20764] - (AB002303) KIAA0305 [Homo sapiens]	Contains protein domain (PF01363) - FYVE zinc finger	struct	29331822, 264908, 264907, 264591, 85658542, 264763, 21906785, 35695917, 264638, 264486
2124	83005951 (4247, 4248)	Novel Protein sim. GBank gij5689455 [dbj]BAA83011.1] - (AB028982) KIAA1059 protein [Homo sapiens]	Contains protein domain (PF00801) - PKD domain	transport	284583
2125	95354041 (4249, 4250)	Novel Protein sim. GBank gij728831 [spj]P39188 [ALU1_HUMAN - III] ALU SUBFAMILY J WARNING ENTRY III]		UNCLASSIFIED	284259, 264509, 264907, 264511, 85658542, 264763, 21906785, 35695917, 264638, 264486
2126	95084231 (4251, 4252)	Novel Protein sim. GBank gij4539264 [lemb]CAB39853.1] - (AL049495) conserved hypothetical protein [Schizosaccharomyces pombe]		UNCLASSIFIED	284488, 264489, 29331827, 35696052, 264905, 264509, 264908, 264909, 264510, 265009, 264591, 264592, 264593, 33657402, 264594, 264595, 264598, 264758, 264601, 264803, 265018, 264804, 264605, 264760, 264681, 264762, 264883, 264764, 264684, 264288, 264685, 264689, 60170815, 33657023, 33657109, 55810764, 264635, 264636, 264637, 264638, 264639, 83373044, 264594, 264566

2127	81118652 (4253, 4254)	Novel Protein sim. GBank glj4868435[gb]AAD31315.1[AF143236] apoptosis related protein APR-2 [Homo sapiens]			35696286, 29331826, 35696052, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 265006, 264511, 264512, 265007, 265009, 264910, 264758, 265011, 264600, 264601, 264604, 264762, 264763, 264766, 264687, 264768, 264769, 264689, 35695917, 264690, 264691, 264692, 264693, 264629, 18108374, 35695855, 264632, 264634, 264635, 264636, 264637, 264638, 264639, 18108385, 264563, 264564, 264567, 56182575, 35696288, 56182181, 29331824, 60432289, 35698052, 264905, 264907, 66712502, 264908, 264909, 264510, 264512, 265009, 264910, 264591, 55612038, 265018, 264764, 264288, 264369, 264687, 264768, 55811957, 264692, 18108388, 264628, 264632, 264634, 264635, 264637, 56182323, 264639, 18108384, 18108388, 264563, 264567
2128	87414262 (4255, 4256)				UNCLASSIFIED
2129	95102069 (4257, 4258)				UNCLASSIFIED
2130	95417144 (4259, 4260)	Novel Protein sim. GBank glj2649255 (AE001012) - conserved hypothetical protein [Archaeoglobus fulgidus]			263981
2131	85723065 (4261, 4262)	Novel Protein sim. GBank glj1086886 (U41276) - Similar to potassium channel protein. [Caenorhabditis elegans]	Contains protein domain (PF00805) - Pentapeptide repeats (8 copies)		35696052, 264909, 264768, 35695917
2132	95361096 (4263, 4264)	Novel Protein sim. GBank glj5689373[dbj]BAA82973.1 - (AB028944) KIAA1021 protein [Homo sapiens]	Contains protein domain (PF00122) - E1-E2 ATPase		264488, 22278999, 264259, 29331827, 29331828, 35698052, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 265006, 264511, 265007, 264910, 264591, 264592, 264595, 264758, 21806754, 33109954, 87168474, 265011, 264600, 264601, 264605, 265019, 264760, 18108351, 264681, 264762, 264784, 264288, 264684, 264766, 264688, 264687, 264768, 264769, 264688, 21908769, 264690, 52644150, 264691, 264693, 18108370, 264628, 264629, 18108372, 18108374, 35696423, 35695855, 264631, 264634, 264635, 264636, 264555, 264637, 18108380, 264639, 264558, 56182323, 56528486, 264564, 264565, 264566, 264567

2133	95351539 (4265, 4266)	Novel Protein sim. GBank gij4220489 (AC006069) - hypothetical protein [Arabidopsis thaliana]		UNCLASSIFIED	60424178, 52646365, 52646842, 56994075, 35698286, 22278997, 22278998, 60432049, 56182181, 66714117, 60424269, 28331828, 29331828, 35696052, 264805, 264908, 264907, 66712502, 29331830, 58182435, 265008, 264512, 265008, 60431735, 60433356, 33657402, 55812038, 33109954, 21908754, 55811386, 265010, 264603, 265017, 265016, 265019, 55811150, 18108351, 264682, 264368, 264288, 52644228, 56181562, 21906765, 21906766, 21908767, 21908768, 21908769, 55811957, 35695917, 265020, 265021, 60170815, 33657023, 33657109, 60431528, 18108374, 35696423, 65274781, 35695855, 264634, 60431850, 18108380, 58182323, 83373044, 18108385, 18108387, 60432113, 22278002, 264563, 264568
2134	95412697 (4267, 4268)	Novel Protein sim. GBank gij3875351 emb CA808415 - (Z96047) DY3.6 [Caenorhabditis elegans]			56181686, 35696286, 21908754, 55811386, 265011, 265017, 18108351, 264765, 264768, 264888, 21908768, 35695917, 265020, 33657023, 264628, 35695855, 264632, 264555, 264556, 264557, 264558, 18108382, 22278002
2135	88078813 (4269, 4270)	Novel Protein sim. GBank gij5689559 dbj BAA83083.1 - (AB028034) KIAA1111 protein [Homo sapiens]	Contains protein domain (PF00628) - PHD-finger	UNCLASSIFIED	22278998, 29331828, 35696052, 264908, 264908, 264910, 265009, 264591, 264758, 52646317, 265011, 87188559, 264601, 18108351, 264448, 264883, 264684, 264689, 18108359, 264691, 33657023, 264692, 35695783, 264629, 35695855, 264631, 264635, 264636, 264637, 58182323, 264639, 22278002, 264564
2136	84346478 (4271, 4272)	Novel Protein sim. GBank gij2662167 dbj BAA23715 - (AB007803) KIAA0443 [Homo sapiens]		UNCLASSIFIED	264639
2137	87637718 (4273, 4274)	Novel Protein sim. GBank gij4884110 emb CAB43262.1 - (AL050090) hypothetical protein [Homo sapiens]		UNCLASSIFIED	264559, 264909, 33109954, 284763, 21908768, 60170394, 18108385, 264563
2138	87395446 (4275, 4276)	Novel Protein sim. GBank gij5174779 gb AA0X0698.1 - (U87804) 50 kDa protein [Caulobacter crescentus]		ATPase_associated	264259, 29331828, 35696052, 264909, 265008, 265017, 265018, 18108351, 264288, 21908768, 33657023, 33657109, 264628, 18108374, 35695855, 264634, 264555, 264558, 264557, 264558, 264559
2139	94643882 (4277, 4278)	Novel Protein sim. GBank gij385082 emb CAA77135 - (Y18350) U2 snRNP auxiliary factor, large subunit [Nicotiana glauca]			264905, 264810, 264591, 55812038, 55811386, 85658542, 264760, 18108351, 18108359, 55811957, 265020, 265021, 33657023, 18108384, 55811576, 83373044, 18108385, 56526486, 264482

2140	87845655 (4279, 4280)	Novel Protein sim. GBank gij4417293jgb AAD20418 - (AC007019) unknown protein [Arabidopsis thaliana]		UNCLASSIFIED	264488, 264259, 29331824, 284104, 264109, 264509, 285008, 264759, 285018, 284448, 264288, 21908768, 55811957, 285021, 33657023, 27488265, 35698423, 264636, 264556, 264557, 284559, 264566
2141	78623986 (4281, 4282)			UNCLASSIFIED	265020, 284693
2142	80041222 (4283, 4284)			UNCLASSIFIED	283978
2143	84140051 (4285, 4286)	Novel Protein sim. GBank gij2135786jpr j S53362 - mucin 5AC (clone JER47) - human (fragment)		UNCLASSIFIED	22278997, 29331827, 284907, 265020, 60432113
2144	94320114 (4287, 4288)	Novel Protein sim. GBank gij2078483 (U43200) - antilreeze glycopeptide AFGP polypeptide precursor [Boreogadus saida]		UNCLASSIFIED	65274572, 264259, 29331824, 29331827, 264908, 264908, 264591, 265011, 87188559, 264800, 285019, 264288, 264768, 21908765, 21908767, 55811576, 35698423, 65274791, 222789002
2145	20564305 (4289, 4290)	Novel Protein sim. GBank gij1255871 (U53341) - short region of weak similarity to bovine membrane receptor p63 (PIR:S28503) [Caenorhabditis elegans]		UNCLASSIFIED	283978
2146	87010515 (4291, 4292)	Novel Protein sim. GBank gij3080398 emb CAA18718.1 - (AL022603) putative protein [Arabidopsis thaliana]		UNCLASSIFIED	264907, 284768, 264769, 18108385
2147	80432911 (4293, 4294)	Novel Protein sim. GBank gij728837 sp P39194 ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII		UNCLASSIFIED	264593
2148	80048811 (4295, 4296)	Novel Protein sim. GBank gij119863 sp P20693 FCE2_MOUSE - LOW AFFINITY IMMUNOGLOBULIN EPSILON FC RECEPTOR (LYMPHOCYTE IGE RECEPTOR) (FC-EPSILON-RII) (CD23)	Contains protein domain (PF00059) - Lectin C-type domain	glycoprotein	29331824, 29331828, 35698052, 264758, 87188474, 285018, 52644150, 33657109
2149	87362022 (4297, 4298)	Novel Protein sim. GBank gij5420387 emb CAB46679.1 - (AJ243459) proteophosphoglycan [Leishmania major]		UNCLASSIFIED	22278998, 29331822, 29331824, 29331828, 264764, 264769, 21908766, 284488
2150	94140059 (4299, 4300)	Novel Protein sim. GBank gij5699407 dbj BAA82987.1 - (AB028958) KIAA1035 protein [Homo sapiens]			22278996, 56994075, 22278999, 60432049, 264259, 29331822, 29331824, 29331826, 35698052, 29331828, 284508, 284511, 80433356, 264758, 264596, 33109954, 60174639, 265010, 265011, 87168559, 265017, 285018, 265019, 264448, 264288, 264689, 21908765, 21908768, 21908768, 265020, 60170815, 33657109, 33657182, 33657349, 18108370, 264635, 264557, 80170394, 18108385, 87168519, 222789000
2151	95353241 (4301, 4302)	Novel Protein sim. GBank gij3452473 (AF084205) - serine/threonine protein kinase TAO1 [Rattus norvegicus]		kinase	18108397, 18108398, 265007, 264591, 265011, 18108351, 18108368, 18108374, 18108388
2152	79321640 (4303, 4304)				

2153	88313371 (4305, 4308)	Novel Protein sim. GBank gij4758704[re]NP_004216.1[PMASL - MFH-amplified sequences with leucine-rich tandem repeats 1	Contains protein domain (PF00560) - Leucine Rich Repeat	glycoprotein	264488, 263994, 52646842, 22278996, 22278998, 22278999, 284259, 29331822, 35698052, 284508, 284509, 264905, 264906, 264907, 264908, 264909, 56182435, 284510, 264511, 284512, 284758, 87188474, 87188559, 265017, 285019, 284780, 284288, 284389, 284766, 284687, 284769, 52644229, 21908768, 21908768, 35695917, 33857023, 33857109, 35695855, 264631, 264632, 284635, 264636, 284639, 18108385, 284483, 284584, 284486
2154	87408034 (4307, 4308)	Novel Protein sim. GBank gij225150[pr]j1209265U - chorion protein B11 [Bombyx mori]		UNCLASSIFIED	56994075, 284094, 265009, 265019, 284286, 21908767, 35695917
2155	87424072 (4308, 4310)			UNCLASSIFIED	18108392, 18108398, 22278996, 264259, 29331824, 265008, 265010, 265011, 265017, 265019, 284288, 284686, 265020, 284693, 284628, 56182323
2156	84285205 (4311, 4312)	Novel Protein sim. GBank gij3970966 (AC004974) - spa-1- like; similar to AF026504 (PID:g2555183) [Homo sapiens]			265007, 284684
2157	87316344 (4313, 4314)	Novel Protein sim. GBank gij1076211[pr]j[S50765 - hypothetical protein VSP-3 - Chlamydomonas reinhardtii]		UNCLASSIFIED	264591
2158	88444218 (4315, 4316)	Novel Protein sim. GBank gij4650844[db]j[BAA77027.1] - (AB026180) Ketch motif containing protein [Homo sapiens]	Contains protein domain (PF00551) - BTB/POZ domain		264598
2159	80083729 (4317, 4318)	Novel Protein sim. GBank gij2879825[db]j[BAA24826] - (AB007897) KIAA0437 [Homo sapiens]			29331822, 284112, 265009, 264691, 33857023, 284634
2160	16283674 (4319, 4320)	Novel Protein sim. GBank gij1504006[db]j[BAA13202] - (D88966) similar to human ZFY protein. [Homo sapiens]		UNCLASSIFIED	284634
2161	87739131 (4321, 4322)			UNCLASSIFIED	285008
2162	84319528 (4323, 4324)			UNCLASSIFIED	63274572, 284508, 264905, 264906, 284907, 284908, 52844045, 284908, 265007, 284910, 264591, 284592, 264593, 55812038, 284596, 264758, 265011, 284600, 284762, 284763, 264683, 284764, 284288, 284768, 284686, 284768, 284769, 284689, 285020, 284691, 284628, 284629, 283878, 284632, 284634, 284557, 284638, 284639, 18108385, 284563, 284566, 284587
2163	85417158 (4325, 4326)	Novel Protein sim. GBank gij3876537[emb]j[CAA98270] - (Z73974) cDNA EST yk29115.3 comes from this gene; cDNA EST yk29115.5 comes from this gene [Caenorhabditis elegans]		UNCLASSIFIED	58182575, 22278998, 264093, 264683, 33857023, 65274820, 60432113
2164	80568456 (4327, 4328)		Contains protein domain (PF01006) - Hepatitis C virus non-structural protein NS4a	collagen	284603, 284637, 264555

2165	94329189 (4329, 4330)	Novel Protein sim. GBank gll1086794 (U41107) - No definition line found [Caenorhabditis elegans]		UNCLASSIFIED	56994075, 22278996, 22278997, 22278998, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 264908, 29331830, 56182435, 265009, 21908754, 33657084, 265011, 265019, 264448, 264288, 264369, 21908765, 21908768, 21908769, 265020, 265021, 264691, 264692, 33657023, 65274820, 35695855, 264556, 60170394, 83373044, 60432113, 22278002, 264567
2166	87618934 (4331, 4332)	Novel Protein sim. GBank gll2706522[emb]CAA75816] - (Y15895) ubiquitin activating enzyme [Drosophila melanogaster]	ubiquitin		52845158, 22278994, 22278998, 66714117, 29331828, 52844045, 265018, 265019, 264369, 21908765, 21908767, 21908768, 21908769, 265021, 265022, 264693, 27486282, 35695763, 18108376, 56528488, 87168518, 264567
2167	87716864 (4333, 4334)	Novel Protein sim. GBank gll2224713[db]BAA20840] - (AB002384) KIAA0386 [Homo sapiens]		UNCLASSIFIED	56182575, 35696288, 29331824, 29331828, 29146498, 56182435, 265008, 265009, 264592, 264593, 33657402, 33109954, 265011, 265017, 265018, 18108351, 264369, 21908764, 21908765, 21908768, 29148627, 21908769, 52844150, 33657109, 35698423, 18108381, 18108384, 18108385, 60432113, 264567
2168	86999334 (4335, 4336)	Novel Protein sim. GBank gll4321407[gb]AAD15748] - (AF047690) ATP-binding cassette protein M-ABC1 [Homo sapiens]	Contains protein domain (PF00684) - ABC transporter transmembrane region.	transport	66714117, 29331827, 264907, 264511, 264591, 265018, 264784, 264683, 264768, 264768, 264568
2169	87866937 (4337, 4338)	Novel Protein sim. GBank gll5106521[gb]AAD39741.1[AF105365] K-CI co-transporter KCC4 [Homo sapiens]		UNCLASSIFIED	264629, 264555, 264559
2170	94141033 (4339, 4340)	Novel Protein sim. GBank gll5106521[gb]AAD39741.1[AF105365] K-CI co-transporter KCC4 [Homo sapiens]		UNCLASSIFIED	65274572, 56182575, 22278997, 22278998, 264259, 29331825, 264509, 264908, 56182435, 60433438, 55812038, 264598, 55811386, 265019, 264762, 264763, 264448, 264764, 264694, 264288, 264766, 264685, 56181562, 264689, 55811957, 265020, 264535, 264691, 33657109, 60431528, 18108374, 35698423, 55811578, 65274791, 264634, 264639, 264558, 87168518, 60432113, 264564
2171	80194050 (4341, 4342)			UNCLASSIFIED	264369, 265020, 264558
2172	85452460 (4343, 4344)			UNCLASSIFIED	264259, 264558
2173	87036740 (4345, 4346)	Novel Protein sim. GBank gll4308681[gb]AAD15478] - (AC006930) R33423.1 [Homo sapiens]		UNCLASSIFIED	264369
2174	95003288 (4347, 4348)	Novel Protein sim. GBank gll2493778[sp]Q09456[YO35_CAEEL - PUTATIVE CUTICLE COLLAGEN C09G5.5]			264908, 35695855, 264555, 264557

2175	84325850 (4349, 4350)	Novel Protein sim. GBank g 1263287 (U47855) - fibroin-3 [Araneus diadematus]		UNCLASSIFIED	264488, 35696288, 20281099, 29331828, 60432289, 35696032, 264109, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 265006, 265007, 264512, 265008, 264910, 264591, 264594, 264595, 264596, 264758, 55812038, 285011, 264600, 264603, 264760, 264762, 264448, 264764, 264288, 264766, 264686, 264887, 21808788, 55811957, 35695917, 265020, 265022, 264691, 264692, 33657023, 264693, 264628, 264629, 55811576, 35696423, 65274781, 35695855, 264631, 264632, 264634, 264635, 264636, 264637, 264638, 264639, 264558, 18108385, 60432113, 264583, 264584, 264565, 264566, 264486, 264567
2176	89223392 (4351, 4352)	Novel Protein sim. GBank g 728837 sp P39184 ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII	Contains protein domain (PF00803) - Pentapeptide repeats (8 copies)	oncogene	52644507, 52846842, 22278994, 35696286, 22278996, 22278998, 29331826, 29331827, 35696032, 29331828, 33656970, 29331830, 264810, 33657402, 264758, 52644296, 87168559, 265018, 264688, 21908765, 21908767, 21908768, 35695917, 52644150, 264690, 33657023, 33657109, 52645129, 33657182, 27486261, 27486262, 33657349, 18108376, 18108377, 35695855, 87168518, 60432113, 264404, 22279000, 264488
2177	94128842 (4353, 4354)	Novel Protein sim. GBank g 5454072 re NP_008418.1 pSLU7 - step II splicing factor SLU7		kinase	18108392, 22278997, 22278999, 264093, 33657402, 265019, 264448, 264766, 264688, 21908767, 21908768, 21908789, 265021, 33657023, 18108370, 18108374, 60432113, 22279002
2178	87601557 (4355, 4356)	Novel Protein sim. GBank g 473407 (U08215) - NST-1 [Mus musculus]	Contains protein domain (PF00012) - Hsp70 protein	eph	264488, 22278996, 22278999, 29331824, 29331825, 29331826, 29331827, 52644296, 87168474, 18108370, 35695855, 22279002
2179	87316275 (4357, 4358)			UNCLASSIFIED	60424288, 264760, 264628, 264632

2180	95351397 (4359, 4360)	Novel Protein sim. GBank gi 3122317 sp P90646 KMH_B_DICDI - MYOSIN HEAVY CHAIN KINASE B (MHCK B)	Contains protein domain (PF00400) - kinase WD domain, G-beta repeat	52644507, 22278994, 35696266, 22278997, 22278999, 264259, 52645080, 29331822, 29331824, 29331825, 29331826, 29331828, 33856970, 264508, 264509, 264908, 264909, 29331830, 264909, 264510, 265008, 265007, 33657402, 55812038, 21908754, 87188474, 87188559, 265017, 265018, 265019, 264783, 264882, 264683, 264684, 264288, 264686, 21908765, 21908768, 21908769, 265020, 265021, 265022, 52844150, 33857023, 33657109, 27486265, 33657349, 18108374, 35898423, 35895855, 263981, 80170394, 18108385, 56528488, 87168518, 60432113, 222789000, 264482, 264568, 264587, 264486 29331827, 264369, 18108376, 264564
2181	85764930 (4361, 4362)	Novel Protein sim. GBank gi 3024889 sp Q15542 T2D4_HUMAN - TRANSCRIPTION INITIATION FACTOR TFIIID 100 KD SUBUNIT (TAFII-100) (TAFII100)		kinase
2182	87637731 (4383, 4384)	Novel Protein sim. GBank gi 5420387 emb CAB46678.1 - (AJ243459) proteophosphoglycan [Leishmania major]		UNCLASSIFIED
2183	85460849 (4385, 4386)	Novel Protein sim. GBank gi 3873406 gb AAC77482.1 - (U17129) unknown [Rhodococcus erythropolis]		
2184	87760690 (4367, 4368)	Novel Protein sim. GBank gi 3114713 (AF061346) - Edp1 protein [Mus musculus]		inf
2185	87826483 (4369, 4370)	Novel Protein sim. GBank gi 5106956 gb AAD39906.1 AF11361 - (AF113615) FH1/FH2 domain-containing protein FHOS [Homo sapiens] Novel Protein sim. GBank gi 2864625 emb CAA16972 - (AL021811) putative protein [Arabidopsis thaliana]		ATPase-associated
2186	87739227 (4371, 4372)			
2187	87388173 (4373, 4374)			UNCLASSIFIED
2188	87771708 (4375, 4376)	Novel Protein sim. GBank gi 5107816 gb AAD40129.1 AF14941 - (AF149413) contains similarity to histone deacetylases; Pfam PF00850, Score=13.3, E=5e-10, N=1 [Arabidopsis thaliana]		histone
2189	85693573 (4377, 4378)	Novel Protein sim. GBank gi 3452357 (AF075724) - unknown [Legionella pneumophila]	Contains protein domain (PF01596) - O-methyltransferase	22278996, 264259, 29331826, 21908754, 264369, 264288, 263987

2180	67639187 (4379, 4380)	Novel Protein sim. GBank gij132575[sp]29315[IRINI_RAT - RIBONUCLEASE INHIBITOR]		nucleaseinhib	22278996, 22278999, 29331822, 29331824, 29331826, 265008, 264910, 60170831, 55812038, 52844298, 265010, 265018, 264685, 264886, 58181562, 21908789, 35695917, 265022, 60170394, 22279000, 29331825, 29331826, 29331830, 264510, 264511, 264910, 264593, 264594, 264556, 264559, 264558
2181	95198928 (4381, 4382)	Novel Protein sim. GBank gij5327002[emb]CAB46272.1] - (Y18503) XAP-5-like protein [Homo sapiens]			
2182	11126316 (4383, 4384)	Novel Protein sim. GBank gij462600[sp]P34400[M110 CAEL - MIG-10 PROTEIN]	Contains protein domain (PF00169) - PH domain		
2183	84140073 (4385, 4386)	Novel Protein sim. GBank gij5420389[emb]CAB46680.1] - (AJ243460) proteophosphoglycan [Leishmania major]		UNCLASSIFIED	58181686, 29331825, 29331827, 264508, 264908, 265008, 264592, 60432229, 264288, 264684, 264786, 35695917, 33657023, 60431802, 60431528, 55810784, 55811578, 65274791, 35695855, 60431850, 56182323, 60432113, 264592
2184	21418714 (4387, 4388)	Novel Protein sim. GBank gij2773341 (AF040954) - putative protein phosphatase 1 nuclear targeting subunit [Rattus norvegicus]			
2185	88083023 (4389, 4390)	Novel Protein sim. GBank gij2832763[emb]CAA15685.1] - (AL009191) /prediction=(method:: /prediction=(method:: /match=(desc:: /match=(desc:: /molli=(desc: [Drosophila melanogaster]		UNCLASSIFIED	22278996, 22278998, 35696052, 265008, 21908754, 265017, 35695917, 265021, 265022, 35695855
2186	95081631 (4391, 4392)	Novel Protein sim. GBank gij5262487[emb]CAB45699.1] - (AL080076) hypothetical protein [Homo sapiens]		collagen	56182575, 35696286, 22278997, 22278999, 264259, 29331822, 66714117, 60432289, 29331827, 35696052, 29331828, 264508, 52844045, 56182435, 264510, 265007, 265008, 265009, 60433438, 55812038, 265010, 265011, 264448, 264288, 264686, 264687, 52844229, 21908785, 21908788, 21908789, 35695917, 265022, 264691, 33657023, 264693, 18108370, 18108376, 35698423, 55811576, 65274791, 35695855, 264638, 56182323, 18108385
2187	95073813 (4393, 4394)	Novel Protein sim. GBank gij4929587[gb]AAD34044.1[AF15180 - (AF151807) CGI-49 protein [Homo sapiens]			264788, 264789, 21908785, 21908786, 21908787, 29148827, 55811957, 35696286, 265020, 22278998, 265021, 264259, 33657023, 264693, 29331824, 35696052, 29331828, 18108370, 35695855, 264113, 265008, 264910, 60432229, 56182323, 33657402, 264758, 83373044, 21908754, 265018, 265019, 22279002, 264482, 264448, 264865, 264288, 264369
2188	88060914 (4395, 4396)	Novel Protein sim. GBank gij3548787 (AC005622) - R30953_1 [Homo sapiens]		UNCLASSIFIED	

2199	88054355 (4397, 4398)	Novel Protein sim. GBank gij2739372 (AC002505) - hypothetical protein [Arabidopsis thaliana]			264105, 264110, 264112, 264688, 55811957, 33857023, 264692, 263987, 20281071, 56526486
2200	87405385 (4398, 4400)	Novel Protein sim. GBank gij3043634[dbj BAA25481] - (AB011127) KIAA0555 protein [Homo sapiens]	struct		29331824, 264763, 264768
2201	94316872 (4401, 4402)	Novel Protein sim. GBank gij3913470[sp O57314 DH8X_ANAPL - PUTATIVE STEROID DEHYDROGENASE SPM2]	Contains protein domain (PF00106) - short chain dehydrogenase		29331824, 35696052, 264905, 264907, 33857402, 55811386, 265017, 265018, 265019, 264288, 21906788, 35695917, 265020, 265022, 33657023, 33657108, 27486261, 18108370, 35696423, 35695855, 264555, 264556, 83373044, 87188518, 60432113
2202	91672385 (4403, 4404)	Novel Protein sim. GBank gij5262665[emb CAB45767.1] - (AL080186) hypothetical protein [Homo sapiens]	UNCLASSIFIED		264489, 264259, 29331824, 60432289, 35696052, 264905, 264909, 264592, 265017, 265018, 265019, 18108351, 264762, 284448, 264369, 264288, 264766, 21906785, 21906766, 264690, 264691, 264692, 33857109, 264634, 264636, 264555, 264639, 264558, 264559, 83373044, 18108385, 264404, 22279002, 264482
2203	87761832 (4405, 4406)	Novel Protein sim. GBank gij172845[sp P46829 RB25_RABIT - RAS-RELATED PROTEIN RAB-25]	Contains protein domain (PF00071) - Ras family	glycoprotein	52646365, 56994075, 264259, 29331822, 29331826, 29331827, 29331828, 284910, 265010, 265011, 87168559, 265018, 265019, 264805, 264288, 21906789, 35695917, 33857023, 264692, 33857109, 35695763, 18108376, 264638, 22279000, 264566, 264567
2204	88088671 (4407, 4408)	Novel Protein sim. GBank gij121036[sp P28348 GBT3_RAT GUANINE NUCLEOTIDE-BINDING PROTEIN G(T), ALPHA-3 SUBUNIT (GUSTDUCIN ALPHA-3 CHAIN)]	Contains protein domain (PF00503) - G-protein alpha subunit	UNCLASSIFIED	
2205	94147569 (4409, 4410)	Novel Protein sim. GBank gij4569480[dbj BAA76768.1] - (AB023141) KIAA0924 protein [Homo sapiens]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	dna_rna_bind	18108394, 18108397, 56182575, 60432049, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 264906, 265007, 265008, 265009, 80432229, 265010, 265011, 265018, 264683, 264288, 264369, 264686, 21906766, 21908768, 21908769, 264690, 264691, 264693, 18108368, 55811576, 65274791, 264634, 18108381, 18108384, 60432113, 22278002, 264563, 264566, 264591
2206	20820008 (4411, 4412)		UNCLASSIFIED		
2207	87787870 (4413, 4414)	Novel Protein sim. GBank gij4557753[rel NP_000372.1 pMID1 - midline 1 protein]	Contains protein domain (PF00622) - SPRY domain		29331822, 56182181, 29331827, 35696052, 52644045, 265008, 265019, 58181582, 55811957, 265021, 33657023, 35695763, 35695855, 60170394, 60432113, 264568
2208	86100830 (4415, 4416)				264908, 265019, 18108351, 21908769
2209	87800420 (4417, 4418)	Novel Protein sim. GBank gij3986746 (AF105228) - tuftelin [Bos taurus]	struct		264112, 265009, 264691, 18108385, 18108374, 264634, 20281166

2210	57152407 (4419, 4420)	Novel Protein sim. GBank gi 728837 sp P39194 ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII		kinase	264603
2211	87341720 (4421, 4422)	Novel Protein sim. GBank gi 728837 sp P39194 ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII		oncogene	264665, 264686, 18108365, 22278002, 264482
2212	91223924 (4423, 4424)	Novel Protein sim. GBank gi 3776027 emb CAA09214 - (AJ010475) RNA helicase [Arabidopsis thaliana]	Contains protein domain (PF00270) - DEAD/DEAH box helicase	helicase	22278995, 22278997, 22278999, 264092, 264094, 29331822, 66714117, 29331828, 29331828, 264807, 52644045, 265008, 60170831, 21908754, 87168559, 265017, 265019, 18108351, 264683, 18108354, 264369, 264786, 264687, 52644228, 21908765, 21908766, 21908767, 21908768, 265021, 33657109, 18108370, 18108374, 264638, 56182323, 18108384, 18108387, 87168518, 264585
2213	91219309 (4425, 4426)	Novel Protein sim. GBank gi 5420387 emb CAB46679.1 - (AJ243459) proteophosphoglycan [Leishmania major]			56182575, 22278998, 22278997, 35686052, 264905, 66712502, 264908, 264828, 56182435, 264112, 265008, 60431735, 60433438, 21908754, 265010, 265011, 265017, 265018, 265019, 18108351, 264765, 21908765, 21908768, 21908769, 265020, 265021, 264693, 264629, 263974, 263978, 18108370, 55811576, 264556, 264637, 264558, 83373044, 22278002, 264482, 264483

2214	95361453 (4427, 4428)	Novel Protein sim. GBank gl14504325[INP_000173.1]pHADH - hydroxyacyl- Coenzyme A dehydrogenase/3-ketoacyl-Coenzyme A thiolase/enoyl-Coenzyme A hydratase (trifunctional protein), alpha su	Contains protein domain (PF00725) - 3-hydroxyacyl-CoA dehydrogenase	264488, 52644507, 18108394, 56182575, 22278994, 22278995, 35698286, 56994075, 22278997, 22278998, 22278999, 264490, 60432049, 264259, 52645080, 29331822, 29147820, 29331824, 66714117, 29331825, 60432289, 29331826, 29331827, 35698052, 29331828, 20281100, 264509, 264907, 66712502, 264908, 29331830, 52644045, 56182435, 264510, 285006, 264511, 264512, 265007, 265008, 265009, 60170831, 264593, 60433358, 60433438, 33109934, 33857084, 52644298, 87168474, 265010, 265011, 87168559, 264601, 265017, 265018, 265019, 18108351, 264448, 264682, 264763, 264288, 264887, 52644229, 264889, 21908765, 21908766, 21908767, 21908768, 21908769, 35695917, 265020, 265021, 265022, 264532, 60170815, 264690, 52644150, 264691, 33857023, 264692, 18108364, 33857109, 33857182, 27488262, 27488264, 27488285, 35695763, 18108370, 264629, 60431528, 18108374, 18108376, 55810764, 35698423, 35695955, 264634, 264636, 52644332, 264638, 264558, 60170394, 18108381, 56182323, 83373044, 18108385, 18108387, 18108388, 56526486, 87168518, 60432113, 22278002, 264482, 264564, 264565, 264566, 264909, 265008, 264555, 264558, 87168518
2215	95419208 (4428, 4430)	Novel Protein sim. GBank gl11947160 (AF000298) - weak similarity to collagens; glycine- and proline-rich [Caenorhabditis elegans]		
2216	87614046 (4431, 4432)	Novel Protein sim. GBank gl11572802 (U70854) - similar to Enterococcus faecalis TRAB (GI:388268) [Caenorhabditis elegans]	UNCLASSIFIED	264693
2217	80589404 (4433, 4434)	Novel Protein sim. GBank gl15031707[INP_005503.1]pGARP - glycoprotein A repetitions predominant	glycoprotein	264288, 33657109, 264556
2218	85518254 (4435, 4436)	Novel Protein sim. GBank gl13878636[emb]CAA88953] - (Z49128) similar to cAMP-dependent protein kinase; cDNA EST EMBL: T00719 comes from this gene; cDNA EST yk465d8.3 comes from this gene; cDNA EST yk465d8.5 comes from this gene; cDNA EST yk492/4.3 comes from this gene; cDNA EST y...	Contains protein domain (PF00560) - Leucine Rich Repeat	35696423, 264563
2219	87614048 (4437, 4438)	Novel Protein sim. GBank gl11572802 (U70854) - similar to Enterococcus faecalis TRAB (GI:388268) [Caenorhabditis elegans]	Contains protein domain (PF01963) - TraB family	264692, 264693, 264688, 264689, 264693, 18108370, 18108378

2220	95354185 (4439, 4440)	Novel Protein sim. GBank g 4507281 ref NP_003145.1 pSTAT - stattherin			264488, 18108394, 18108395, 35698286, 264259, 264097, 60432289, 264509, 264905, 264908, 264907, 29331830, 264908, 264909, 264510, 264511, 265007, 264512, 264910, 265009, 264593, 264594, 60433358, 264595, 55812038, 264758, 85658542, 285010, 264601, 264603, 265019, 264605, 264760, 264762, 264448, 264764, 264388, 264768, 18108357, 264768, 264687, 18108358, 264769, 55811957, 264690, 264691, 33657023, 264692, 18108362, 18108368, 264628, 264629, 18108374, 263978, 264634, 264635, 264636, 264637, 264638, 18108385, 264483, 264568, 264488, 264587
2221	88060927 (4441, 4442)	Novel Protein sim. GBank g 3549154 (AC005625) - R27328_1 [Homo sapiens]			
2222	84425892 (4443, 4444)			UNCLASSIFIED	
2223	95091849 (4445, 4446)			UNCLASSIFIED	264908, 265020, 35695855 265010, 264685, 264690, 264693, 264628, 263974, 263976, 55811576, 264555, 264638, 83373044, 264483
2224	87388515 (4447, 4448)	Novel Protein sim. GBank g 3876005 emb CAA84789 - (Z35719) cDNA EST EMBL:D87419 comes from this gene; cDNA EST EMBL:C13853 comes from this gene; cDNA EST EMBL:C11578 comes from this gene; cDNA EST YK234a7.3 comes from this gene; cDNA EST YK234a7.5 comes from this gene; cDNA ES...	Contains protein domain (PF01958) - Domain of unknown function	UNCLASSIFIED	264259, 264509, 56182435, 265008, 265008, 265009, 264757, 21908754, 18108351, 264693, 18108374, 18108385
2225	85749484 (4449, 4450)	Novel Protein sim. GBank g 1255847 (U53338) - C05E11.1 gene product [Caenorhabditis elegans]		transport	22278994, 22278895, 22278999, 52844045, 264600, 265019, 21908785, 21908789
2226	86978953 (4451, 4452)	Novel Protein sim. GBank g 4826524 emb CAB42852.1 - (AL049848) hypothetical protein [Homo sapiens]			264259, 29331822, 29331824, 29331825, 29331827, 264508, 264908, 265007, 264691, 264634, 264488
2227	87721135 (4453, 4454)			UNCLASSIFIED	22278999, 265008, 265008, 18108354, 29148629, 29148784, 27486281, 18108374, 264637, 18108384
2228	91227337 (4455, 4456)	Novel Protein sim. GBank g 606976 (U16800) - ribonucleoprotein [Xenopus laevis]	Contains protein domain (PF00076) - RNA recognition motif (e.a.a. RRM, RBD, or RNP domain)	- dna_rna_bind	264091, 264092, 264094, 29331822, 29331825, 68714117, 264693, 263972, 264639, 83373044, 264583
2229	88060931 (4457, 4458)	Novel Protein sim. GBank g 3549155 (AC005625) - R27328_2 [Homo sapiens]		UNCLASSIFIED	

2230	95342915 (4459, 4460)	Novel Protein sim. GBank gij226154iprj11412350A - DNA polymerase [Human adenovirus type 2]		UNCLASSIFIED	264488, 264768, 2644507, 264769, 21906765, 21906768, 21906767, 21906769, 22278995, 35695917, 22278996, 22278997, 22278998, 22278999, 265021, 264259, 2645129, 29331827, 264508, 264509, 264907, 18108370, 18108374, 35696423, 35695855, 265007, 264910, 264555, 33657402, 21906754, 18108387, 265010, 265018, 265019, 264760, 264288, 264587, 264563
2231	98060937 (4461, 4462)	Novel Protein sim. GBank gij3549154 (AC005625) - R27328_1 [Homo sapiens]		UNCLASSIFIED	
2232	87762581 (4463, 4464)	Novel Protein sim. GBank gij5281316ipb AAD41478.1 AF133124 transcription factor IIC63 [Homo sapiens]		transcription factor	18108394, 56162575, 22278995, 35696288, 22278997, 22278999, 264259, 29331827, 35696052, 264907, 56182435, 285006, 265007, 265008, 264910, 264758, 55812038, 264603, 265018, 265019, 18108351, 264682, 264764, 264683, 264369, 264288, 264686, 264687, 264689, 21906765, 21906768, 21906767, 21906769, 26148829, 35695917, 264690, 52844150, 264691, 33657023, 264693, 18108370, 18108374, 55811576, 35695855, 264639, 18108385, 264564, 264908, 33657402, 265018, 264288, 264686, 265020, 264635, 18108385
2233	87755292 (4465, 4466)	Novel Protein sim. GBank gij4249733ipb AAD137801 - (AF109377) IdBp [Mus musculus]		kinase	56182435, 264369, 264688, 21906765, 265020, 264693, 264556, 56528468
2234	87771817 (4467, 4468)	Novel Protein sim. GBank gij1706559ip54352JEAS, DROME - ETHANOLAMINE KINASE (EASILY SHOCKED PROTEIN)		UNCLASSIFIED	22278997, 264563
2235	91012316 (4469, 4470)	Novel Protein sim. GBank gij4972734ipb AAD34782.1 - (AF132174) unknown [Drosophila melanogaster]	WD domain, G-beta repeat		264569, 264687, 264769, 265022, 264259, 60432049, 264691, 29331828, 60432289, 20281149, 264906, 264907, 264511, 265008, 265009, 264634, 264635, 264636, 264555, 264556, 264557, 264558, 60433358, 264595, 264559, 60433438, 60432113, 264781, 264762, 264763, 264784
2236	88003131 (4471, 4472)	Novel Protein sim. GBank gij1082675iprj1853814 - p20 protein - human	Contains protein domain (PF00011) - eph		
2237	91012318 (4473, 4474)	Novel Protein sim. GBank gij4972734ipb AAD34782.1 - (AF132174) unknown [Drosophila melanogaster]	Contains protein domain (PF000400) - WD domain, G-beta repeat	kinase	264488, 264769, 21906765, 21906768, 21906767, 21906768, 21906769, 22278995, 22278996, 22278997, 22278998, 265020, 265021, 22278999, 264259, 29331824, 29331826, 29331827, 35695763, 18108378, 35695855, 265007, 60432229, 33657402, 60433358, 60433438, 83373044, 18108385, 21906754, 18108387, 60432113, 22279000, 265019, 22279002, 264482, 18108351, 264288

2238	84998857 (4475, 4478)			Contains protein domain (PF00286) - Viral coat protein	264509, 264907, 264629, 264634, 264584
2239	87796688 (4477, 4478)				28331825, 265009, 264359, 33657109, 18108370, 18108374, 264557, 264559
2240	84121471 (4479, 4480)	Novel Protein sim. GBank g 2982311 (AF051240) - probable ubiquitin-conjugating enzyme E2 [Picea mariana]		Contains protein domain (PF00179) - Ubiquitin-conjugating enzyme	284488, 65274572, 56182575, 35698286, 22278997, 22278998, 264259, 29331827, 35696052, 264508, 52644045, 58182435, 264511, 265007, 285008, 265009, 80433356, 80433438, 55812038, 21908754, 33657084, 55811388, 265018, 265019, 18108351, 264683, 264288, 264768, 264687, 264688, 264769, 21908765, 21908789, 21908789, 35695917, 265021, 265022, 60170815, 52644150, 33657023, 33657182, 33657349, 35695763, 18108370, 35696423, 35695855, 87188518, 22279000
2241	80091951 (4481, 4482)				284693, 264629
2242	81228075 (4483, 4484)	Novel Protein sim. GBank g 2494312 sp P70541 E2BQ_RAT - TRANSLATION INITIATION FACTOR EIF-2B GAMMA SUBUNIT (EIF-2B GDP-GTP EXCHANGE FACTOR)		UNCLASSIFIED synthase	22278995, 22278998, 22278997, 22278998, 264259, 29331822, 29331824, 29331826, 29331827, 29331828, 264509, 285007, 265008, 284598, 21908754, 265010, 285011, 285017, 265018, 265019, 264448, 264389, 264288, 52644229, 21908765, 21908768, 21908767, 21908768, 21908769, 265020, 265021, 33657109, 27488262, 27488264, 18108374, 35695855, 264634, 264637, 56182323, 63373044, 56526488, 87188518, 264564
2243	78902028 (4485, 4486)			UNCLASSIFIED	265008
2244	85723527 (4487, 4488)	Novel Protein sim. GBank g 2291143 (AF016417) - Similar to BZIP transcription factor [Caenorhabditis elegans]		UNCLASSIFIED	264604
2245	85316545 (4489, 4490)	Novel Protein sim. GBank g 470340 (U00043) - similar to beta-mannosyltransferase [Caenorhabditis elegans]	Contains protein domain (PF00534) - Glycosyl transferases group 1	UNCLASSIFIED	52645156, 22278995, 22278996, 22278997, 22278999, 29331822, 29331824, 29331827, 264907, 264512, 60433438, 264758, 21908754, 265011, 264603, 264764, 264687, 21908767, 21908768, 21908769, 55811957, 265022, 264681, 264629, 35698423, 264638, 18108387, 60432113, 22279000, 22279002, 264588

2248	94848710 (4491, 4492)	Novel Protein sim. GBank gl 4986096 dbj BAA78328.1 - (AB028089) activator of S phase Kinase [Homo sapiens]	Contains protein domain (PF00153) - transport Mitochondrial carrier proteins	65274572, 22278995, 35698288, 22278996, 22278997, 22278999, 284259, 35698052, 264108, 264905, 284907, 265008, 265007, 265008, 60433438, 33108954, 87168559, 265018, 265019, 284288, 21906765, 21906767, 21906768, 21906769, 55811957, 35695917, 265020, 265022, 27486284, 18108370, 18108374, 65274791, 35695855, 60432113
2247	87862542 (4493, 4494)	Novel Protein sim. GBank gl 854065 emb CAA50337 - (X83413) U88 [Human herpesvirus 6]	UNCLASSIFIED	52645158, 52646385, 52645080, 35698052, 33696970, 52848317, 33657084, 265017, 21906768, 21906769, 35695917, 33657109, 52845129, 33657182, 27486261, 27486282, 33657349, 27486285, 18108387
2248	95412996 (4495, 4496)	Novel Protein sim. GBank gl 4758502 ref NP_004123.1 pHABP - hyaluronan-binding protein 2	Contains protein domain (PF00089) - Trypsin	264488, 264259, 284907, 29331830, 284909, 265007, 265009, 264595, 21906754, 65274444, 284603, 265019, 264762, 284448, 284288, 264689, 21906768, 55811957, 265021, 284691, 18108374, 284634, 284635, 284636, 264555, 284838, 264557, 264558, 284559, 18108383, 83373044, 18108385, 284488
2249	9485662 (4497, 4498)	Novel Protein sim. GBank gl 403846 (AF107772) - TcSTII [Trypanosoma cruzi]	Contains protein domain (PF00515) - eph TPR Domain	264768, 264628, 284636, 264637
2250	79827508 (4499, 4500)	Novel Protein sim. GBank gl 3738140 emb CAA21241 - (AL031852) valyl-tRNA synthetase, mitochondrial precursor [Schizosaccharomyces pombe]	UNCLASSIFIED	264908, 18108374
2251	87385863 (4501, 4502)	Novel Protein sim. GBank gl 3218467 emb CAA07090.1 - (AJ006529) putative phosphatase [Gallus gallus]	UNCLASSIFIED	284259, 35698052, 284508, 58182435, 265009, 264592, 284593, 284760, 264448, 264884, 284288, 284690, 264628, 55811578, 264555, 264556, 264557, 264558, 264559, 264568
2252	87735867 (4503, 4504)	Novel Protein sim. GBank gl 4929325 gb AAD33953.1 AF14531 - (AF145316) vacuolar proton pump delta polypeptide [Homo sapiens]	Contains protein domain (PF01813) - synthase ATP synthase subunit D	264092, 264094, 264259, 29331822, 68714117, 29331828, 284102, 284103, 284104, 264105, 284109, 284112, 264511, 265007, 60433356, 265010, 18108351, 21906767, 21906768, 284691, 263974, 263977, 284486, 284567
2253	91010703 (4505, 4506)		UNCLASSIFIED	65274572, 265019

2254	95320031 (4507, 4508)	Novel Protein sim. GBank gi 4502847 ref NP_001271.1 pCIRB - cold inducible RNA-binding protein	Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	dna_rna_bind	264569, 18108394, 18108398, 56182575, 56994075, 35698286, 22278998, 284084, 60432049, 284259, 29331822, 29331824, 29331825, 29331826, 60432289, 29331827, 29331828, 35698052, 264108, 264508, 264509, 284808, 284907, 29331830, 66712502, 264908, 284909, 284510, 265006, 284511, 265007, 265008, 265009, 60170831, 60432229, 60433356, 60433438, 264758, 85698542, 265010, 265011, 87168559, 265017, 265018, 265019, 264448, 264764, 284288, 264369, 264766, 264686, 284768, 284769, 21906765, 21906767, 55811857, 264691, 33657023, 264692, 18108382, 65274620, 263969, 264628, 18108370, 60431528, 263972, 264629, 18108372, 18108377, 18108379, 55811578, 35698423, 35698555, 264630, 264634, 264635, 264636, 264556, 263981, 264638, 56182323, 60170394, 264558, 18108381, 18108382, 83373044, 18108385, 87168516, 60432113, 22278002, 264482, 264584, 264565, 264486, 264567, 18108391
2255	91010548 (4509, 4510)	Novel Protein sim. GBank gi 5541865 emb CA851072.1 - (AL088858) hypothetical protein [Homo sapiens]	Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	struct	65274572, 56182575, 22278997, 22278999, 284259, 29331822, 29331825, 29331826, 29331827, 29331828, 264508, 264905, 284906, 264907, 66712502, 264908, 56182435, 264510, 264511, 265008, 264593, 264595, 21906754, 33109954, 87168474, 265011, 265017, 265019, 284682, 264764, 284389, 264288, 264766, 264685, 264686, 284768, 21906765, 21906766, 21906768, 21906769, 265020, 60170615, 52644150, 264690, 264692, 264693, 33657109, 33657349, 264632, 264636, 52644332, 56182323, 22279000, 22279002 264768, 264689, 18108374
2256	87020531 (4511, 4512)	Novel Protein sim. GBank gi 3327174 dbj BA31655 - (AB014580) KIAA0680 protein [Homo sapiens]		UNCLASSIFIED	
2257	80088235 (4513, 4514)			UNCLASSIFIED	22278996, 22278998, 264681, 21906765, 21906768, 264567
2258	88090516 (4515, 4516)	Novel Protein sim. GBank gi 3025446 (AC004528) - R32184 2 [Homo sapiens]	Contains protein domain (PF00080) - Ligand-gated ion channel	misc_channel	264908, 264592, 264764

2259	95364155 (4517, 4518)	Novel Protein sim. GBank gij4884140[emb]CAB43278.1] - (AL050110) hypothetical protein [Homo sapiens]		UNCLASSIFIED	18108398, 65274572, 58182575, 22278997, 22278998, 264259, 29331822, 29331827, 284805, 88712502, 284908, 264909, 58182435, 265007, 265008, 60432229, 33857084, 87188559, 18108351, 284448, 284683, 284288, 264369, 58181582, 265021, 60170615, 264890, 33857109, 60431528, 18108374, 52644332, 58182323, 18108385, 22278000, 22278002, 264482
2260	88084119 (4518, 4520)	Novel Protein sim. GBank gij3080863 (AC004614) - similar. to T-aponin proteins AB008086 (PID:g2529225) [Homo sapiens]	Contains protein domain (PF00090) - Thrombospondin type 1 domain	oxidase	58182575, 22278999, 29331822, 29331825, 80432289, 29331827, 35896052, 284508, 88712502, 52844045, 58182435, 285008, 265008, 265009, 60433356, 55812038, 285010, 285017, 285019, 284288, 284389, 21906765, 21908767, 55811957, 35695917, 52644150, 33857023, 33857109, 55811578, 65274791, 58182323
2261	88074157 (4521, 4522)	Novel Protein sim. GBank gij3334528[emb]CAA18138] - (AL021308) predicted using FGENEH [Homo sapiens]		UNCLASSIFIED	22278994, 22278997, 284907, 284828, 52644150, 18108361, 284693, 18108374
2262	9163282 (4523, 4524)	Novel Protein sim. GBank gij487759gb[AAD31421.1]AF12444 - (AF124440) MAGE tumor antigen D1 [Homo sapiens]		UNCLASSIFIED	284686, 284488, 284768, 284769, 284691, 284508, 284905, 284509, 284908, 284907, 284908, 284909, 35895855, 284510, 284511, 284512, 285007, 285009, 284638, 284639, 284757, 284758, 18108385, 285011, 284760, 284564, 284565, 284764, 264568, 284486, 284766
2263	87602495 (4525, 4526)	Novel Protein sim. GBank gij3341697 (AC003872) - hypothetical protein [Arabidopsis thaliana]			
2264	87756525 (4527, 4528)	Novel Protein sim. GBank gij1657601 (U66220) - unknown [Nannocystis exedens]			
2265	86918663 (4529, 4530)	Novel Protein sim. GBank gij477072[pil]A48018 - mucin 7 precursor, salivary - human	Contains protein domain (PF00086) - Zinc finger, C2H2 type	UNCLASSIFIED	22278995, 22278997, 22278999, 264259, 285008, 265007, 285009, 60433436, 21906754, 285010, 285011, 285017, 264448, 284683, 284288, 264689, 21908785, 21906768, 35695917, 285021, 18108374, 284638, 22278000, 22278002, 264568, 284487
2266	87773458 (4531, 4532)	Novel Protein sim. GBank gij3150479 (AF067212) - partial CDS [Caenorhabditis elegans]	Contains protein domain (PF01305) - Ribosomal protein L15 amino terminal region	ribosomalprot	

2267	87395638 (4533, 4534)	Novel Protein sim. GBank gij3560228[emb]CAA20697.1]- (AL031530) hypothetical protein [Schizosaccharomyces pombe]	UNCLASSIFIED	3569286, 264259, 29331824, 29331825, 3569052, 29331828, 264905, 264509, 264907, 264908, 264909, 264512, 265009, 264910, 264593, 33657402, 265010, 265018, 264782, 264448, 264288, 264389, 264768, 52644229, 35695917, 264691, 33657023, 18108382, 33657109, 35698423, 264634, 18108381, 87168518, 264566
2268	85693667 (4535, 4536)	Novel Protein sim. GBank gij728832[sp]P39189[ALU2_HUMAN - IIII SUBFAMILY SB WARNING ENTRY IIII protein - fruit fly (Drosophila melanogaster)]	cadherin	264488, 264259, 264509, 264595, 265010, 265017, 264766, 18108385, 264488
2269	88177877 (4537, 4538)	Novel Protein sim. GBank gij103418[pi]S17885 - Tcd37 protein	UNCLASSIFIED	56182575, 60432049, 265007, 265009, 264591, 87168559, 264605, 18108351, 21908784, 265020, 264628, 60431528, 264638, 18108385, 18108387, 60432113 264763
2270	80410327 (4539, 4540)			
2271	91010382 (4541, 4542)		cyl450	264809, 56182435, 265008, 55812038, 55811957, 33657023, 264693, 33657109, 55810784, 55811576, 58182323
2272	84208220 (4543, 4544)		UNCLASSIFIED	264905, 264908
2273	85014271 (4545, 4546)	Novel Protein sim. GBank gij4178370 (AC005058) - similar to calcium-independent phospholipase A2; similar to AC004392 (PID:g3367519) [Homo sapiens]	Contains protein domain (PF00462) - Glutaredoxin	52645156, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 35696052, 264909, 265008, 264593, 60433438, 21908754, 265018, 264689, 21908785, 21908768, 21908767, 21908769, 265021, 265022, 60170615, 264691, 33657023, 264693, 33657109, 27488264, 18108376, 35696423, 35695855, 264630, 52644332, 264558, 58182323, 22279002
2274	91840217 (4547, 4548)	Novel Protein sim. GBank gij1480112[emb]CAA67981]- (X99842) HP1-BP74 protein [Mus musculus]	Contains protein domain (PF00538) - linker histone H1 and H5 family	52645156, 22278997, 22278999, 52645080, 29331824, 29331825, 29331826, 29331827, 29331828, 264905, 264908, 52644045, 264511, 265008, 265009, 60170831, 264591, 21908754, 33109954, 265011, 265018, 18108351, 264448, 264288, 264684, 264768, 21908765, 21908766, 21908767, 21908768, 52644150, 264693, 18108384, 35695783, 18108374, 35698423, 264634, 264557, 264638, 52644332, 83373044, 18108385, 58528486, 87168518, 22279002
2275	88082501 (4549, 4550)	Novel Protein sim. GBank gij3165406 (AC004755) - fos37502_2 [Homo sapiens]	Contains protein domain (PF00122) - E1-E2 ATPase	
2276	11287447 (4551, 4552)		transport	
			UNCLASSIFIED	264555, 264558

2277	88084123 (4553, 4554)	Novel Protein sim. GBank gij2880079 (AC004142) - similar to murine leucine-rich repeat protein; possible role in neural development by protein-protein interactions; 93% similarity to D49802 (PID:g1389908) [Homo sapiens]	Contains protein domain (PF00560) - Leucine Rich Repeat	glycoprotein	2227899, 35696052, 265008, 265019, 284369, 265020, 265022, 55810784, 284404, 22279002
2278	94133079 (4555, 4556)	Novel Protein sim. GBank gij2618702 (AC002510) - unknown protein [Arabidopsis thaliana]		UNCLASSIFIED	18108394, 22278997, 29331826, 60433356, 60433438, 21906754, 265018, 33657023, 284639, 83373044, 284565
2279	80419375 (4557, 4558)	Novel Protein sim. GBank gij119714sp P13983 EXTN_TOBAC - EXTENSIN PRECURSOR (CELL WALL HYDROXYPROLINE-RICH GLYCOPROTEIN)		UNCLASSIFIED	284766, 284565
2280	94239723 (4559, 4560)				
2281	95283048 (4561, 4562)	Novel Protein sim. GBank gij4240299 dbj BAA74928.1 - (AB020712) KIAA0805 protein [Homo sapiens]	Contains protein domain (PF00400) - WD domain, G-beta repeat	transport	265008, 33109954, 285010, 285019, 265020, 264092, 284259, 29331822, 29331824, 29331826, 35696052, 264107, 264908, 284909, 52844045, 265008, 33657402, 60433356, 284758, 265011, 285019, 284681, 284683, 284684, 284686, 21908765, 21908767, 21908768, 21908769, 60170615, 264690, 52844150, 18108382, 264692, 18108388, 18108374, 283978, 264631, 18108381, 284559, 18108385, 56526486, 22279000, 284568, 284567
2282	87602829 (4563, 4564)	Novel Protein sim. GBank gij1537070 (U63840) - nucleoporin p54 [Rattus norvegicus]		UNCLASSIFIED	284488, 264259, 29331822, 29331824, 29331827, 29331828, 29331830, 33657402, 60433438, 87168474, 285019, 18108351, 21908767, 21908769, 55811957, 33657023, 52845129, 33657109, 33657182, 27486282, 283972, 55811578, 87168518, 20281189, 80424179, 56182575, 22278994, 35696286, 22278997, 22278999, 29331822, 29331824, 56182181, 29331825, 29331827, 35696052, 28146499, 284905, 66712502, 284908, 285007, 265009, 80432229, 284593, 60431735, 60433356, 33109954, 33657084, 55811388, 87168474, 285010, 265011, 285018, 285019, 55811150, 284683, 284369, 284288, 284688, 21906765, 21908767, 21908768, 28148627, 21908769, 55811957, 265020, 265022, 33657182, 27486281, 18108370, 284628, 18108374, 55810784, 18108379, 55811576, 35696423, 35695855, 284630, 60431850, 283981, 18108382, 83373044, 18108385, 18108387, 60432113, 22279000, 284482, 284567
2283	95382386 (4565, 4566)	Novel Protein sim. GBank gij2495729sp Q82556 Y281_HUMAN - HYPOTHETICAL PROTEIN KIAA0281 (HA6725)		UNCLASSIFIED	

2284	95414955 (4567, 4568)	Novel Protein sim. GBank gi 2498797 sp Q84311 PNAD_MOUSE - PROTEIN N- TERMINAL ASPARAGINE AMIDOHYDROLASE (PROTEIN NH2-TERMINAL ASPARAGINE DEAMIDASE) (NTN- AMIDASE) (PNAD) (PROTEIN NH2-TERMINAL ASPARAGINE AMIDOHYDROLASE) (PNAAL)			60424178, 52644507, 18108394, 52846842, 22278994, 35688288, 22278998, 22278997, 22278999, 284259, 60432049, 28331822, 28331824, 28331825, 60432289, 28331826, 28331827, 35898052, 28331830, 52844045, 56182435, 33657402, 60433438, 33109954, 21908754, 85658542, 87188559, 265018, 265019, 59811150, 264682, 264389, 21908765, 21908768, 21908767, 21908768, 21908769, 55811857, 35695917, 265020, 265021, 60170615, 33657023, 33857182, 27488262, 27488284, 27488265, 18108376, 55810764, 35698423, 35695855, 60431850, 87188518, 60432113, 284482, 264564
2285	87761484 (4569, 4570)	Novel Protein sim. GBank gi 3342234 (U83909) - nuclear antigen EBNA-1 [Cercopithecine herpesvirus 15]	collagen		35696052, 284905, 264907, 264808, 264909, 264512, 285009, 264910, 264595, 264780, 18108351, 264682, 264763, 264685, 264768, 264686, 264768, 264693, 264629, 35695855, 264631, 264634
2286	87737825 (4571, 4572)	Novel Protein sim. GBank gi 3873414 (U00043) - similar to D. melanogaster trithorax protein [Caenorhabditis elegans]	kinase		35688288, 58182435, 60170831, 264591, 60432229, 264592, 264593, 264594, 264595, 55812038, 264596, 87188474, 35695917, 264692, 55811576, 264555, 264557
2287	82986696 (4573, 4574)	Novel Protein sim. GBank gi 830905 sp j S42731 - collagen alpha 1 chain - sea urchin (Hemicentrotus pulcherrimus) (fragment)	UNCLASSIFIED	Contains protein domain (PF01391) - Collagen triple helix repeat (20 copies)	264682
2288	94133083 (4575, 4576)	Novel Protein sim. GBank gi 728832 sp P38189 ALU2_HUMAN - III ALU SUBFAMILY SB WARNING ENTRY IIII	lm7		22278995, 56894075, 22278997, 22278999, 264259, 60432288, 264508, 264512, 265008, 33657402, 265017, 265018, 265019, 18108351, 264448, 21908765, 21908766, 21908767, 21908768, 33657023, 264557, 22278900, 22279002
2289	88084133 (4577, 4578)	Novel Protein sim. GBank gi 2887497 (AC004144) - R34001_1 [Homo sapiens]	kinase	Contains protein domain (PF00400) - WD domain, G-beta repeat	
2290	88084137 (4579, 4580)	Novel Protein sim. GBank gi 2887497 (AC004144) - R34001_1 [Homo sapiens]	UNCLASSIFIED	Contains protein domain (PF00400) - WD domain, G-beta repeat	264683
2291	84285281 (4581, 4582)	Novel Protein sim. GBank gi 3253120 (AC005175) - R31449_3 [Homo sapiens]	struc		08394, 264807, 265006, 265008, 33109954, 52846317, 265010, 18108351, 264681, 264686, 264692, 18108370, 18108374, 18108385

2292	94328834 (4583, 4584)	Novel Protein sim. GBank gij4803872[emb]CA842843.11 - (AJ133789) nuclear transport receptor [Homo sapiens]		UNCLASSIFIED	56182575, 35696286, 56994075, 29331824, 29331825, 35696052, 56182435, 60433438, 55612038, 33109954, 87168474, 87168559, 265018, 18108351, 264763, 264448, 264389, 264288, 56181562, 264769, 21906765, 21906768, 21906767, 21906769, 265021, 265022, 33657023, 264693, 65274620, 33657109, 27488264, 264629, 55810764, 55811576, 35695855, 56182323, 56526486, 87168518, 22278000, 264567
2293	87759213 (4585, 4586)	Novel Protein sim. GBank gij3252981 (AF068921) - Ras-binding protein SUR-8 [Mus musculus]	Contains protein domain (PF00560) - struct Leucine Rich Repeat		264488, 18108397, 35696286, 264092, 264259, 29331822, 29331826, 264908, 264908, 264511, 264512, 265009, 264910, 18108351, 264764, 264369, 264288, 264885, 264766, 265020, 265022, 264534, 35696423, 264631, 264637, 18108381, 56182323, 264639, 18108385, 264404, 264563, 264565, 22278997, 22278998, 22278999, 264259, 29331822, 60432289, 29331828, 35696052, 265018, 264684, 264288, 264686, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 265021, 264692, 33657109, 18108376, 35696423, 35695855, 264634, 22278900, 22278902, 264563, 264488, 22278996, 60432289, 264682, 264683, 264689, 18108374
2294	86693580 (4587, 4588)	Novel Protein sim. GBank gij2062680 (U88964) - HEM45 [Homo sapiens]	Contains protein domain (PF00929) - nuclease Exonuclease		263974, 263978, 22278999, 264259, 29331822, 29331824, 29331826, 29331828, 35696052, 264907, 264908, 52644045, 56182435, 264112, 265006, 265007, 264910, 265009, 60433358, 33657402, 264595, 55812038, 21906754, 265011, 265018, 265019, 264448, 264764, 264288, 264766, 21906765, 21906767, 21906768, 21906769, 55811957, 265020, 265021, 264690, 264691, 33657023, 264692, 264693, 33657109, 55811576, 56182323, 60170394, 83373044, 18108385, 56526486, 264584, 264486
2295	95312200 (4589, 4590)			UNCLASSIFIED	
2296	80030781 (4591, 4592)	Novel Protein sim. GBank gij5689501[dbj]BAA83034.11 - (AB029005) KIAA1082 protein [Homo sapiens]		transcript factor	
2297	84321251 (4593, 4594)				

2298	95312207 (4595, 4596)	Novel Protein sim. GBank gi 2875051 emb CAB02849 - (Z81050) predicted using GeneFinder; similar to collagen; cDNA EST EMBL:D65584 comes from this gene; cDNA EST EMBL:D69046 comes from this gene; cDNA EST yk366b12.3 comes from this gene; cDNA EST yk366b12.5 comes from this gene ...		collagen	60424179, 56161688, 22278995, 35698288, 22278996, 22278998, 22278999, 264480, 264259, 29331822, 29331824, 66714117, 60424288, 35698052, 29331828, 66712502, 56182435, 264510, 265006, 60433436, 21906754, 33108954, 55811386, 265010, 265018, 55811150, 264762, 18108351, 264862, 264863, 264288, 264684, 264888, 264888, 56181562, 264689, 21906768, 21906767, 29148629, 55811857, 29148784, 35695917, 265020, 18108382, 33657023, 18108384, 33657109, 60431602, 18108370, 60431528, 18108374, 55810784, 35698423, 35695955, 264630, 264634, 60431650, 18108380, 56182323, 264558, 83373044, 18108385, 60432113, 22278000, 264482, 264567, 264488
2299	80193720 (4597, 4598)			UNCLASSIFIED	264369
2300	84124346 (4599, 4600)	Novel Protein sim. GBank gi 2443886 (AC002294) - Unknown protein [Arabidopsis thaliana]			264488, 22278996, 22278998, 264259, 29331824, 66714117, 35698052, 264509, 264905, 264908, 264907, 264908, 264909, 265008, 264910, 265009, 264758, 265010, 87168559, 264600, 265018, 264760, 264762, 18108351, 264764, 264766, 264788, 264789, 21906768, 21906767, 35695917, 285021, 264691, 33657023, 35695763, 18108370, 18108374, 35698423, 35695855, 264631, 264636, 264638, 18108385, 22278002, 264563
2301	91235725 (4601, 4602)	Novel Protein sim. GBank gi 2143637 pir I84505 - calcium-dependent actin-binding protein - rat		struct	264908, 264758, 265017, 21906765, 83373044, 264563
2302	88084141 (4603, 4604)	Novel Protein sim. GBank gi 2867497 (AC004144) - R34001_1 [Homo sapiens]		UNCLASSIFIED	52644045, 265019, 264288, 33657023, 18108370, 18108385
2303	94141439 (4605, 4606)	Novel Protein sim. GBank gi 4884194 emb CAB43220.1 - (AL049946) hypothetical protein [Homo sapiens]		struct	264259, 60432049, 264907, 264909, 264910, 60432229, 33657402, 265011, 285018, 264762, 264448, 264789, 264637, 264638, 83373044, 264488
2304	94840434 (4607, 4608)	Novel Protein sim. GBank gi 2494182 p Q10005 YRY1_CAEEL - HYPOTHETICAL 39.9 KD PROTEIN T15H9.1 IN CHROMOSOME II PRECURSOR		UNCLASSIFIED	264259, 29331824, 21906767, 33657182, 33657349
2305	90935911 (4609, 4610)	Novel Protein sim. GBank gi 4972686 gb AA034738.1 - (AF132150) unknown [Drosophila melanogaster]			65274572, 22278998, 264908, 265008, 21906769, 264691, 264488

2308	95334940 (4611, 4612)	Novel Protein sim. GBank gi4929565 p AA034043.1 AF15180 - (AF151806) CGI-48 protein [Homo sapiens]	Contains protein domain (PF00400) - WD domain, G-beta repeat	kinasereceptor	264488, 22278995, 22278998, 35686286, 22278997, 22278998, 22278998, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 35698052, 284508, 264905, 264907, 29331830, 264908, 264909, 264510, 285006, 264511, 265007, 264512, 265008, 265009, 33657402, 21905754, 85658542, 285010, 285011, 284601, 265017, 285018, 264604, 285019, 18108351, 264448, 264288, 264788, 264789, 21906785, 21906768, 21906767, 21906768, 29148829, 29148784, 35695917, 265020, 265021, 265022, 33657023, 264692, 18108370, 18108374, 18108376, 35698423, 35695955, 264630, 264634, 264635, 264636, 264637, 264638, 264639, 18108382, 18108385, 18108387, 264563, 264566, 264488
2307	79415283 (4613, 4614)	Novel Protein sim. GBank		UNCLASSIFIED	264828
2308	87608409 (4615, 4616)	gi4758732 p NP_004522.1 pMOCS - molybdenum cofactor synthesis 2		synthase	35696286, 264259, 29331822, 29331824, 264112, 264512, 264757, 21908754, 264288, 264690, 27486264, 264631, 264634, 264404
2309	95357218 (4617, 4618)	Novel Protein sim. GBank gi3878059 emb CAB17070 - (Z99942) cDNA EST EMBL:D73444 comes from this gene; cDNA EST EMBL:D70905 comes from this gene; cDNA EST EMBL:D72208 comes from this gene; cDNA EST EMBL:D75030 comes from this gene; cDNA EST EMBL:D72944 comes from this gene; cDN...		glycoprotein	18108397, 22278996, 22278997, 22278998, 22278999, 60432049, 29331822, 29331828, 60432289, 66712502, 60432228, 60433356, 60433438, 65274444, 285010, 264600, 264681, 264448, 264683, 264288, 21908768, 21908768, 285020, 264691, 264692, 264693, 65274620, 65274791
2310	79601668 (4619, 4620)	Novel Protein sim. GBank gi2137337 p J148281 - gene	Contains protein domain (PF00013) - KH domain	UNCLASSIFIED	264508
2311	87721189 (4621, 4622)	mC9P protein - mouse		transcriptfactor	18108397, 58182575, 22278998, 56984075, 264259, 29331824, 29331827, 264508, 264907, 58182435, 264510, 264511, 265008, 284512, 285007, 265008, 265009, 60433438, 33109954, 265010, 265011, 264603, 265017, 18108351, 284782, 264683, 264288, 264369, 264688, 33657023, 20281149, 20281069, 264628, 263972, 55811576, 35698423, 20281071, 264632, 264636, 18108385, 18108387, 87188518, 22279000, 264563, 264488

2312	87549881 (4623, 4624)	Novel Protein sim. GBank gj2911264 (AC002550) - Unknown gene product [Homo sapiens]		55182575, 56994075, 35696286, 22278996, 22278997, 22278999, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 52644045, 60432229, 60433356, 55812038, 33109954, 21906754, 87168474, 265018, 18108351, 264288, 52644229, 21906765, 21906768, 21906767, 21906768, 21906769, 265022, 60170615, 33657023, 27486281, 27486264, 35696423, 35695855, 18108385, 22278000, 22278002 263981
2313	80042533 (4625, 4626)	Novel Protein sim. GBank gj3043626(jb) BAA254771-(AB011123) KIAA0551 protein [Homo sapiens]		
2314	94313401 (4627, 4628)	Novel Protein sim. GBank gj5596714(jemb) CAB51401.11-(AL035398) dJ786117.2 (CGI-51) [Homo sapiens]	UNCLASSIFIED	52644507, 52646365, 52646842, 22278996, 22278997, 22278999, 264259, 52645080, 29331822, 29331824, 60432289, 29331826, 29331827, 29331828, 35696052, 35696970, 52644045, 265008, 264593, 60433356, 60433438, 264756, 33109954, 265010, 265017, 265018, 265019, 264288, 264369, 21906765, 21906768, 21906768, 35695917, 52644150, 33657023, 33657109, 52645129, 33657349, 35695763, 18108374, 35696423, 35695855, 52644332, 22278000, 22278002, 264583, 264587
2315	80430119 (4629, 4630)		UNCLASSIFIED	264905, 264906, 264767, 264768, 264693, 55811576, 264635, 56182323, 18108385, 52644507, 52645156, 52646365, 22278994, 22278995, 22278996, 56994075, 35696286, 22278997, 22278998, 22278999, 264259, 29331822, 52645080, 29331824, 29331825, 29331826, 29331827, 35696052, 29331828, 264908, 264807, 29331830, 52644045, 56182435, 264511, 265007, 265008, 265009, 60170831, 60433438, 21906754, 52646317, 33109954, 33657084, 52644286, 87168474, 265010, 87168559, 265017, 265018, 265019, 264681, 264763, 264448, 264683, 264368, 52644229, 21906764, 264688, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 52644150, 33657023, 18108362, 52645129, 33657182, 33657349, 35695763, 18108370, 18108376, 35696423, 35695855, 264631, 264556, 52644332, 83373044, 18108385, 18108387, 87168518, 60432113, 22278000, 264586, 264587
2316	94312191 (4631, 4632)	Novel Protein sim. GBank gj5531827(jb) AAD44488.11-(AF078856) p47 [Homo sapiens]	Contains protein domain (PF00789) - glycoprotein UBX domain	

2317	87020571 (4633, 4634)		UNCLASSIFIED	22278998, 60432049, 264810, 60432228, 264686, 264687, 264688, 264689, 264558, 18108385
2318	79959879 (4635, 4636)		UNCLASSIFIED	265006, 264910
2319	95101781 (4637, 4638)	Novel Protein sim. GBank gij5282613[emb CAB45746.1] - (AL080155) hypothetical protein [Homo sapiens]		264486, 264569, 18108396, 52646365, 22278994, 22278995, 22278996, 56994075, 35686286, 22278997, 22278998, 264259, 52645080, 26331825, 26331826, 26331827, 26331828, 26331830, 56182435, 60170831, 60432229, 60431735, 33857402, 21908754, 52644298, 87168474, 265011, 87168559, 265017, 265018, 265019, 18108351, 264448, 18108354, 264288, 264389, 52644229, 21908764, 21906765, 21906766, 21908767, 21908768, 21906768, 265021, 265022, 52644150, 33857023, 52645129, 33657109, 27486284, 33657349, 35695763, 18108370, 18108376, 18108378, 35698423, 264558, 83373044, 18108385, 56526486, 87168518, 264584, 264565, 264566
2320	91622426 (4639, 4640)	Novel Protein sim. GBank gij728837[sp P39194 ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII]	kinase	22278994, 60432049, 60432289, 29331827, 264511, 265008, 52646317, 265017, 265019, 21908765, 18108372, 18108387, 22279002, 264486, 264687, 18108394, 264689, 21908765, 18108397, 18108398, 21908767, 21908768, 65274791, 22278995, 35695855, 22278998, 265021, 265022, 264510, 265008, 264511, 264512, 265008, 60170815, 264555, 264636, 264556, 18108361, 264259, 60432229, 33657023, 264557, 264558, 264693, 60433356, 264559, 60433438, 29331824, 18108365, 18108348, 18108384, 29331825, 18108385, 33109954, 29331827, 56526486, 29146499, 265011, 60432113, 265017, 265018, 264508, 264563, 264482, 264509, 18108351, 264448, 264907, 264682, 18108370, 264683, 264908, 264288, 264909, 18108354, 264486, 264587
2321	94320377 (4641, 4642)	Novel Protein sim. GBank gij3873837[emb CAB02700] - (Z81029) Similarity to S.pombe hypothetical protein C1D4.09C (SW:Q10154); cDNA EST EMBL: T00543 comes from this gene; cDNA EST EMBL: T01062 comes from this gene; cDNA EST EMBL: T01321 comes from this gene; cDNA EST EMBL: T02288 com...	UNCLASSIFIED	
2322	87803165 (4643, 4644)	Novel Protein sim. GBank gij567857[emb CAB51685.1] - (AL109630) BACR7A4.y [Drosophila melanogaster]	Contains protein domain (PF00106) - dehydrogenase short chain dehydrogenase	22278996, 264907, 264511, 264757, 18108351, 264768, 264638

2323	94840445 (4645, 4646)	Novel Protein sim. GBank gi 2494162 sp Q10005 YRY1_CAEEL - HYPOTHETICAL 39.9 KD PROTEIN T15H9.1 IN CHROMOSOME II PRECURSOR	Contains protein domain (PF00226) - eph DnaJ domain	22278994, 22278995, 22278997, 60432049, 264259, 29331822, 33656970, 264509, 56182435, 284511, 265008, 60433356, 60433438, 55812038, 33109954, 21908754, 85858542, 87168474, 285011, 87188559, 285017, 285019, 264760, 264681, 18108351, 264369, 264288, 18108355, 264687, 264688, 21908765, 21908767, 21908768, 55811957, 35695917, 265021, 33657023, 18108362, 27486262, 55811576, 264631, 284555, 83373044, 87168518, 60432113, 22278002 264592, 284593, 265020
2324	86633807 (4647, 4648)	Novel Protein sim. GBank gi 5419865 emb CAB46377.1 - (AL098732) hypothetical protein [Homo sapiens]	ATPase-associated	265006, 264759, 35695855, 56182323
2325	88165074 (4649, 4650)	Novel Protein sim. GBank gi 231885 sp P29981 CP4C_BLADI - CYTOCHROME P450 4C1 (CYP1VC1)	Contains protein domain (PF00067) - Cytochrome P450	
2326	84390962 (4651, 4652)	Novel Protein sim. GBank gi 231885 sp P29981 CP4C_BLADI - CYTOCHROME P450 4C1 (CYP1VC1)	UNCLASSIFIED	
2327	88081848 (4653, 4654)	Novel Protein sim. GBank gi 4240227 db BAA74892.1 - (AB020676) KIAA0869 protein [Homo sapiens]	UNCLASSIFIED	264259, 264508, 264905, 264906, 264907, 264908, 265007, 264512, 264910, 264758, 265010, 264766, 264768, 264769, 33657023, 264693, 264628, 264631, 264634, 264638, 264639, 284486
2328	83388428 (4655, 4656)	Novel Protein sim. GBank gi 1169343 sp P42209 DIF8_MOUSE - DIFF8 PROTEIN	UNCLASSIFIED	60433438, 264595, 265017, 264766, 264692, 264629, 264635, 264636, 264638, 56182323, 60432113, 264566
2329	87604478 (4657, 4658)	Novel Protein sim. GBank gi 1169343 sp P42209 DIF8_MOUSE - DIFF8 PROTEIN	UNCLASSIFIED	265017, 264685, 60432113, 264088
2330	87335398 (4659, 4660)	Novel Protein sim. GBank gi 5679136 gb AA046874.1 AF16093 - (AF160934) BcDNA.LD14189 [Drosophila melanogaster]	transport	265009
2331	86990463 (4661, 4662)	Novel Protein sim. GBank gi 5679136 gb AA046874.1 AF16093 - (AF160934) BcDNA.LD14189 [Drosophila melanogaster]	ATPase-associated	35696286, 22278998, 29331824, 60424269, 265008, 265008, 265018, 264448, 264764, 21908765, 35695917, 35695855, 264638, 22279000, 284566
2332	87784182 (4663, 4664)	Novel Protein sim. GBank gi 2104452 emb CAB08779 - (Z95397) unknown [Schizosaccharomyces pombe]	UNCLASSIFIED	56182575, 56994075, 29331826, 29331828, 264107, 33657402, 87188559, 264683, 35695917, 265021, 33657023, 263976
2333	88206958 (4665, 4666)	Novel Protein sim. GBank gi 3879985 emb CAA92691.1 - (Z68318) cDNA EST CEMSD62F comes from this gene; cDNA EST EMBL: C07830 comes from this gene; cDNA EST EMBL: C09493 comes from this gene; cDNA EST YK41568.3 comes from this gene; cDNA EST yk41568.5 comes from this gene; cDNA EST...	UNCLASSIFIED	
2334	94319788 (4667, 4668)	Novel Protein sim. GBank gi 4966270 gb AA052261.2 - (U97002) similar to acyl-CoA dehydrogenases and epoxide hydrolases; Pfam domain PF00441 (Acyl-CoA_dh). Score=57.4, E-value=1.7e-16, N=2; contains similarity to Pfam domain PF00702 (Hydrolase). Score=57.4, E- value=1e-13, N=1 IC...	Contains protein domain (PF00441) - Acyl-CoA dehydrogenase	56182575, 29331825, 21908768, 264636, 83373044

2335	80046103 (4669, 4670)	Novel Protein sim. GBank gl 3283350 (AF062378) - calmodulin-binding protein SH-A1 [Mus musculus]	Contains protein domain (PF00612) - struct	18108351, 21906769, 264555
2336	95196121 (4671, 4672)	Novel Protein sim. GBank gl 1929055(jemb CAA72805) - (Y12090) putative 3,4-dihydroxy-2-butanone kinase [Lycopersicon esculentum]	Kinase	264907, 35695917, 18108379
2337	95345810 (4673, 4674)	Novel Protein sim. GBank gl 4485063(jemb CAB39181.1) - (Z85988) dJ108K1.1.3 (similar to yeast suppressor protein SRP40) [Homo sapiens]	UNCLASSIFIED	35696286, 22278999, 56182181, 29331825, 60424289, 56182435, 33657402, 55812038, 55811388, 265017, 265018, 285019, 21908768, 35695917, 284891, 33657023, 33657109, 263972, 35698423, 35695855, 60432113
2338	87634045 (4675, 4676)	Novel Protein sim. GBank gl 2224689(jb BAA20829) - (AB002372) KIAA0374 [Homo sapiens]	Contains protein domain (PF00323) - UNCLASSIFIED	18108394, 29331822, 66714117, 60432289, 284908, 56182435, 265009, 60433438, 284596, 265010, 265019, 18108354, 264288, 264389, 55811957, 265021, 33657023, 263976, 55811576, 264632, 56182323, 284639
2339	85663319 (4677, 4678)	Novel Protein sim. GBank gl 3873550(jemb CAA22127) - (AL033534) serine-rich protein [Schizosaccharomyces pombe]	UNCLASSIFIED	35698286, 264592, 264369, 264691, 264558
2340	90937716 (4679, 4680)			65274572, 22278994, 35696286, 22278997, 22278999, 264259, 29331822, 60432289, 29331826, 29331830, 265009, 33657402, 33109954, 265017, 265018, 264768, 264685, 21908769, 35695917, 284891, 264692, 35698423, 87188518, 22278000
2341	87775281 (4681, 4682)	Novel Protein sim. GBank gl 3874563(jemb CAB02797) - (Z81042) similar to Yeast hypothetical protein YEF6 like; cDNA EST yk208h5.3 comes from this gene; cDNA EST yk208h5.5 comes from this gene; cDNA EST yk303h1.3 comes from this gene; cDNA EST yk303h1.5 comes from this gene; cDNA ...	Contains protein domain (PF00400) - Kinase	264259, 264908, 264909, 264682, 22279000
2342	95334968 (4683, 4684)		WD domain, G-beta repeat	264489, 65274572, 22278995, 22278996, 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 29331825, 66714117, 29331826, 60432289, 35698052, 284905, 264509, 29331830, 285008, 60170831, 60432229, 60433356, 87168474, 285017, 265018, 265019, 284448, 264369, 284288, 21906765, 21908768, 55811957, 35695917, 265020, 265022, 52844150, 33657023, 65274620, 33657109, 18108370, 18108376, 55810784, 35698423, 55811578, 284556, 284558, 18108385, 60432113, 284563, 264564, 264565, 284568, 284587, 284907, 264512, 265011, 264683
2343	87775448 (4685, 4686)	Novel Protein sim. GBank gl 4929741(jb AAD34131.1)AF15189 - (AF151894) CGI-136 protein [Homo sapiens]	UNCLASSIFIED	
2344	78953198 (4687, 4688)	Novel Protein sim. GBank gl 2506307(isp P13944)CAIC_CHICK - COLLAGEN ALPHA 1(XII) CHAIN PRECURSOR (FIBROCHIMERIN)	UNCLASSIFIED	284758
2345	94319789 (4689, 4690)		Contains protein domain (PF00092) - collagen	284488, 264259, 68712502, 264759, 83373044, 264566

2348	94131820 (4691, 4692)	Novel Protein sim. GBank gll1255411 (U53153) - one short region of weak similarity to S. cerevisiae protease A inhibitor 3 (SP:P01094) and another short region of weak similarity to S. cerevisiae glucose repression mediator protein (SP:P14922) [Caenorhabditis elegans]	Contains protein domain (PF00515) - TPR Domain	proteaseinhib	35696286, 22278998, 264259, 35698052, 29331828, 33857402, 80433356, 33108954, 87168559, 264603, 265019, 18108351, 264881, 264685, 21908766, 285021, 33857109, 55811576, 35695845, 264637, 52644332, 264557, 83373044, 22279000, 22279002
2347	95330367 (4693, 4694)				22278997, 264511, 264683, 264684, 284788, 264687, 264688, 264691, 264692, 55811576
2348	95196133 (4695, 4696)	Novel Protein sim. GBank gll1929056[emb]CAAT2805] - (Y12080) putative 2,4-dihydroxy-2-butanone kinase [Lycopersicon esculentum]		kinase	18108394, 35698286, 264259, 35698052, 264508, 264509, 264905, 264906, 264807, 264908, 264909, 264510, 264511, 265008, 265007, 264512, 265008, 265009, 264910, 264591, 264592, 264593, 264594, 264757, 264595, 264596, 264758, 265011, 264601, 264762, 18108351, 264764, 284288, 284768, 264769, 264889, 35695917, 264893, 264828, 18108370, 264629, 18108374, 35698423, 264631, 264635, 264636, 264637, 264638, 264639, 83373044, 18108385, 264567, 264486
2349	87776502 (4697, 4698)	Novel Protein sim. GBank gll4984106[emb]CAB43254.1] - (AL050082) hypothetical protein [Homo sapiens]			35698052, 29146499, 264909, 284369
2350	88260594 (4699, 4700)				22278998, 22278999, 264259, 29331822, 29331824, 29331825, 29331827, 29331828, 33109954, 21906754, 265010, 87168559, 285018, 265019, 284781, 284681, 284288, 18108357, 21906766, 21906767, 264691, 284692, 35695855, 87168518, 22279000, 22279002, 264482
2351	86988042 (4701, 4702)	Novel Protein sim. GBank gll728832[sp]P39189[ALU2_HUMAN - III] ALU SUBFAMILY SB WARNING ENTRY IIII		kinase	56182575, 284909, 265006, 264558
2352	87337186 (4703, 4704)	Novel Protein sim. GBank gll731637[sp]P38760[YHH5_YEAST - HYPOTHETICAL 75.9 KD PROTEIN IN SPO13-ARG4 INTERGENIC REGION]	Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	UNCLASSIFIED	264259, 264448
2353	91638784 (4705, 4706)	Novel Protein sim. GBank gll1346945[sp]P48809[R827_DROME - HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN 27C (HNRNP 48) (HRP48.1)]	Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	dna_ma_bind	29331826, 55812038, 265019, 264692, 284636
2354	87337199 (4707, 4708)	Novel Protein sim. GBank gll731637[sp]P38760[YHH5_YEAST - HYPOTHETICAL 75.9 KD PROTEIN IN SPO13-ARG4 INTERGENIC REGION]	Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	UNCLASSIFIED	29331824, 264908, 265006, 265008

2355	91638786 (4709, 4710)	Novel Protein sim. GBank gij4938503[emb]CAB43661.1] - (AL078465) hnRNP-like protein [Arabidopsis thaliana]	Contains protein domain (PF00076) - RNA recognition motif, (a.k.a. RRM, RBD, or RNP domain)	dna_rna_bind	56994075, 22278996, 35696286, 22278999, 284259, 29331825, 29331826, 29331828, 29148498, 284905, 264908, 265008, 284758, 87168474, 265010, 265017, 284687, 21908765, 21908767, 21908769, 264691, 284692, 263967, 18108370, 87168518, 22279000
2356	95327688 (4711, 4712)	Novel Protein sim. GBank gij5138920[gb]AAD40377.1] - (AF092135) PTD014 [Homo sapiens]			52644507, 22278995, 35696286, 22278996, 22278997, 22278998, 22278999, 284259, 29331824, 86714117, 29331825, 60432289, 35696052, 29331828, 264908, 66712502, 284512, 285007, 285008, 60170831, 60432229, 60433358, 60433438, 284758, 52646317, 33109954, 21908754, 55611386, 87168474, 265017, 265018, 264605, 265019, 284681, 284682, 284448, 284389, 284288, 284686, 284768, 21908765, 21908768, 21908767, 21908768, 21908769, 285021, 60170815, 33657109, 27486284, 35695763, 55810784, 18108379, 35898423, 55811576, 35895855, 60170394, 56182323, 83373044, 18108385, 56526486, 264404, 60432113, 22279000, 284482, 284563, 264566, 284486, 264567
2357	87775458 (4713, 4714)	Novel Protein sim. GBank gij4929741[gb]AAD34131.1[AF15189 - (AF151894) CGI-136 protein [Homo sapiens]		UNCLASSIFIED	284488, 284769, 18108394, 284259, 29331822, 18108370, 18108374, 264510, 285017, 284482, 284563, 284762, 284565, 284566, 284389, 18108354
2358	87777078 (4715, 4716)	Novel Protein sim. GBank gij4218005 (AC006135) - putative vicilin storage protein (globulin-like) [Arabidopsis thaliana]		UNCLASSIFIED	22278997, 22278999, 284509, 284905, 284592, 18108351, 284681, 284682, 284769, 32833986, 18108374, 284556, 18108385, 284482
2359	87755859 (4717, 4718)	Novel Protein sim. GBank gij1066830 (U41264) - coded for by C. elegans cDNA yk2018.5; coded for by C. elegans cDNA yk44g1.5; coded for by C. elegans cDNA yk12b7.5; coded for by C. elegans cDNA yk36g6.5; coded for by C. elegans cDNA yk2018.5; coded for by C. elegans cDNA yk18g12....		UNCLASSIFIED	35696286, 22278998, 264905, 284511, 285007, 265008, 60433438, 284288, 284686, 21908769, 285020, 284692, 35895855, 284558, 56526488, 284563
2360	80046125 (4719, 4720)	Novel Protein sim. GBank gij3881545[emb]CAA937791 - (Z69904) cDNA EST yk428d5.3 comes from this gene; cDNA EST yk428d5.5 comes from this gene [Caenorhabditis elegans]		UNCLASSIFIED	22278997, 29331826, 263981, 22279000
2361	84232181 (4721, 4722)	Novel Protein sim. GBank gij746487 (U23514) - No definition line found [Caenorhabditis elegans]			22278995, 22278999, 284512, 285009, 284757, 21908765, 65274620, 18108370, 60431528, 18108374, 284635, 60170394, 284482

2362	911721193 (4723, 4724)	Novel Protein sim. GBank gij1171083isp19706MYSB_ACACA - MYOSIN HEAVY CHAIN IB (MYOSIN HEAVY CHAIN II)		UNCLASSIFIED	22278998, 264259, 29331822, 29331824, 60432289, 264509, 264512, 60432228, 60433356, 264448, 264682, 264683, 264369, 21908785, 21908788, 21908789, 60432113, 22279000, 22278002 264907, 264629, 264635
2363	95006635 (4725, 4726)	Novel Protein sim. GBank gij854065[emb]CAA583371 - (X83413) U88 [Human herpesvirus 6]		UNCLASSIFIED	
2364	94827104 (4727, 4728)	Novel Protein sim. GBank gij5639830[gbl]AAD45886.1JAF14601 - (AF146018) hydroxypyruvate reductase [Homo sapiens]	Contains protein domain (PF00389) - O-isomer specific 2-hydroxyacid dehydrogenases	reductase	264488, 18108394, 264887, 18108388, 22278996, 58994075, 35896286, 22278997, 22278998, 264259, 66714117, 29331825, 35896052, 264509, 264905, 264908, 264907, 264908, 66712502, 264909, 264511, 285008, 264512, 265007, 265008, 33657402, 264758, 21906754, 87168474, 265010, 87168559, 264603, 265017, 265018, 265019, 264760, 264762, 18108351, 264448, 264764, 264683, 264684, 264288, 18108355, 264768, 18108358, 264689, 18108359, 21906765, 21908766, 21908767, 35895917, 265020, 265021, 265022, 60170815, 52644150, 264691, 33657023, 264692, 18108364, 33657109, 18108368, 18108370, 18108374, 35898423, 35898555, 264635, 264556, 264557, 264639, 60170394, 83373044, 18108383, 18108384, 18108385, 18108388, 58528486, 264482, 264584, 264486 22278996, 22278998, 22278999, 264907, 264909, 264910, 33657402, 264758, 264600, 264766, 264687, 264689, 21908785, 21908767, 21908768, 21908769, 265021, 33657023, 33657109, 83373044, 264566
2365	94140746 (4729, 4730)	Novel Protein sim. GBank gij1840045 (U49082) - transporter protein [Homo sapiens]		transport	
2366	94312388 (4731, 4732)			UNCLASSIFIED	52644507, 264259, 29331824, 29331825, 29331826, 29331827, 29331828, 264907, 29331830, 264809, 264511, 265008, 33657402, 264595, 52646317, 265017, 265018, 265019, 264605, 264685, 264766, 264689, 21908766, 21908769, 35895917, 265020, 265021, 265022, 52644150, 35895855, 52644332, 18108385, 18108387, 264584, 264586 35898286, 21908768, 55810764, 85274781, 264587 264628
2367	94140910 (4733, 4734)	Novel Protein sim. GBank gij1085457 (U40410) - C54G7.4 gene product [Caenorhabditis elegans]	Contains protein domain (PF00400) - WD domain, G-beta repeat	UNCLASSIFIED	
2368	84322190 (4735, 4736)				

2369	94314334 (4737, 4738)	Novel Protein sim. GBank gij5360501[dbj BAA02158.1] - (AB028343) a-helix coiled-coil rod homologue [Homo sapiens]		siut	52644507, 52646842, 35696286, 264092, 264094, 52645080, 35698052, 284107, 29331830, 52644045, 265006, 265007, 265009, 52644298, 52644229, 284688, 21908765, 21908766, 35695917, 265020, 52644150, 263987, 33857109, 27486285, 35695763, 18108370, 263974, 18108374, 18108376, 52644332, 263981, 18108385, 264508, 264908, 264596
2370	79804120 (4739, 4740)			UNCLASSIFIED	264389
2371	57280408 (4741, 4742)			UNCLASSIFIED	263987, 263981
2372	87642413 (4743, 4744)			UNCLASSIFIED	29331828, 265010, 265019, 35695917, 284634, 60432113
2373	87418611 (4745, 4746)	Novel Protein sim. GBank gij4589582[dbj BAA76813.1] - (AB023186) KIAA0989 protein [Homo sapiens]		UNCLASSIFIED	
2374	94123685 (4747, 4748)	Novel Protein sim. GBank gij5105131[dbj BAA80445.1] - (AF000081) 246aa long hypothetical ribonuclease PH [Aeropyrum pernix]	Contains protein domain (PF01138) - 3' exoribonuclease family	UNCLASSIFIED	265008, 265007, 265008, 265009, 265011, 264766, 35695917, 35695855, 263981, 264557, 264565
2375	87731355 (4749, 4750)	Novel Protein sim. GBank gij1351115[sp P47758]SRPB_MOUSE - SIGNAL RECOGNITION PARTICLE RECEPTOR BETA SUBUNIT (SR-BETA)		UNCLASSIFIED	60432049, 29331824, 264807, 52644045, 264512, 60433358, 21908754, 52644298, 87168559, 264448, 21908765, 21908768, 21908769, 33857023, 18108368, 55811576, 52644332
2376	87613744 (4751, 4752)	Novel Protein sim. GBank gij2645435 (AF007780) - CHD3 [Drosophila melanogaster]	Contains protein domain (PF00628) - PHD-finger	ATPase-associated	264259, 29331830, 264909, 264810, 265009, 60433438, 21908754, 265017, 265018, 265019, 264682, 264268, 264685, 21908767, 263972, 35695855, 87168518, 60432113
2377	95319689 (4753, 4754)	Novel Protein sim. GBank gij5257005[gb AAD41239.1] - (AF083249) Rb binding protein homolog [Homo sapiens]	Contains protein domain (PF01388) - ARID DNA binding domain	UNCLASSIFIED	18108394, 65274572, 22278997, 22278999, 264095, 29331822, 29147620, 29331824, 68714117, 29331825, 29331826, 29331828, 33656970, 29146498, 29146499, 284509, 265008, 265007, 265008, 265009, 60170831, 265010, 265011, 265018, 55811150, 18108351, 264764, 264288, 21908767, 21908768, 29148627, 29148629, 265021, 33657023, 33857109, 18108370, 18108374, 18108378, 35698423, 264558, 83373044, 18108385, 18108388, 56528486, 22279000, 22279002, 264563
2378	94137032 (4755, 4756)	Novel Protein sim. GBank gij1072199 (U40942) - No definition line found [Caenorhabditis elegans]		UNCLASSIFIED	65274572, 56182575, 35698286, 264259, 29331822, 29331824, 66714117, 29331825, 29331826, 60432289, 29331827, 58182435, 264510, 265009, 60433358, 87168474, 265011, 265018, 284288, 21908765, 33657023, 264557, 56182323, 83373044, 18108385, 22279002, 264482
2379	65444324 (4757, 4758)	Novel Protein sim. GBank gij3337357 (AC004481) - hypothetical protein [Arabidopsis thaliana]	Contains protein domain (PF00400) - WD domain, G-beta repeat	Kinase	265017, 264288, 21908768

2380	88823062 (4759, 4760)	Novel Protein sim. GBank gij4502939jreijNP_001845.1lpCOL1 - collagen, type XI, alpha 1	Contains protein domain (PF01410) - Fibrillar collagen C-terminal domain	collagen	264908, 264910, 265011
2381	87608241 (4761, 4762)	Novel Protein sim. GBank gij4455609jembjCAB36555] - (AL031846) dJ742C19.5 (novel Chromobox protein) [Homo sapiens]	Contains protein domain (PF00385) - 'chromo' (CHR)romalin Organization Modifier) domain	helicase	56182575, 264091, 264093, 264259, 29331825, 264105, 264908, 60433356, 21908754, 265017, 265019, 264883, 264288, 264685, 264686, 264687, 264691, 264692, 264693, 55811576, 264638, 264567
2382	91225982 (4763, 4764)	Novel Protein sim. GBank gij4325130jgbjAAD17276] - (AF119716) dmi-2 protein [Drosophila melanogaster]	Contains protein domain (PF00828) - PHD-finger	transport	29331824, 60432289, 264905, 264586, 21908754, 264788, 265022, 264893, 263967, 33657109, 264629, 264631, 264558, 83373044, 60432113, 264482
2383	87442841 (4765, 4766)	Novel Protein sim. GBank gij1902982jobjBAA19005] - (D89049) lectin-like oxidized LDL receptor [Bos taurus]	Contains protein domain (PF00059) - Lectin C-type domain	glycoprotein	265009, 21908765, 21908766
2384	95354766 (4767, 4768)	Novel Protein sim. GBank gij2462851 (AF016252) - Spinophilin [Rattus norvegicus]	Contains protein domain (PF00595) - PDZ domain (Also known as DHR or GLGF).	struct	264488, 52644507, 52645156, 52648365, 35898286, 22278999, 52845080, 29331824, 29331826, 35696052, 29331828, 264908, 264828, 52644045, 265006, 265008, 265009, 33109954, 33657084, 52644296, 265011, 265017, 265018, 264683, 52644229, 21908765, 21908767, 21908768, 265020, 52644150, 33657023, 264693, 65274820, 52645129, 33657109, 33657182, 27488261, 27488282, 27486284, 33657349, 27488285, 35695763, 18108374, 35695855, 264634, 264555, 264556, 264557, 52644332, 264558, 264559, 18108385, 22279000, 22279002
2385	95419485 (4769, 4770)			UNCLASSIFIED	264488, 52644507, 52645156, 264887, 52646385, 22278995, 22278996, 22278997, 22278999, 264259, 52645080, 29331822, 29331826, 35896052, 52644045, 265008, 265007, 265008, 265009, 264910, 60432229, 60433356, 52646317, 21908754, 265018, 264448, 264683, 264686, 264687, 264689, 21908765, 21908766, 21908767, 21908769, 55811957, 265021, 265022, 264690, 264691, 264692, 65274820, 33657109, 18108370, 264631, 52644332, 22279000, 22278902, 264563, 264565, 264567
2386	94742848 (4771, 4772)	Novel Protein sim. GBank gij4929899jgbjAAD34110.1jAF15187 - (AF151873) CGI-115 protein [Homo sapiens]		glycoprotein	264488, 22278995, 22278996, 22278997, 264259, 29148498, 264112, 264511, 60170831, 60432229, 264595, 60433438, 87168474, 67168559, 264682, 21908765, 21908766, 21908767, 21908769, 28148629, 35695917, 265021, 264690, 33657109, 264628, 18108376, 83373044, 60432113, 22279000, 264564, 264568, 264487

2387	14987990 (4773, 4774)			UNCLASSIFIED	264534
2388	11424804 (4775, 4776)			UNCLASSIFIED	264595
2389	95310650 (4777, 4778)	Novel Protein sim. GBank gll4758059[ref]NP_004372.1 pCREB - cAMP responsive element binding protein-like 1	Contains protein domain (PF00170) - bZIP transcription factor	dna_rna_bind	264488, 22278998, 22278999, 264509, 264905, 264908, 264907, 264908, 264909, 265008, 264511, 264512, 264910, 264591, 21908754, 264601, 264604, 264761, 18108351, 264764, 264288, 264768, 264768, 264769, 21908765, 21908768, 264692, 264693, 35698423, 264635, 264638, 264555, 83373044, 22278000, 264488
2390	94320912 (4778, 4780)	Novel Protein sim. GBank gll1844239[dbj]BAA122231 - (D84103) mitochondrial DNA polymerase gamma [Homo sapiens]	Contains protein domain (PF00476) - DNA polymerase family A	polymerase	52644507, 58182575, 22278995, 35698288, 22278996, 22278997, 22278999, 29331822, 29331825, 29331826, 35698052, 284905, 52644045, 265009, 264758, 264759, 33109954, 52644298, 85659542, 265011, 265017, 265018, 264605, 52644229, 21908765, 21908767, 21908768, 21908769, 35695917, 52644150, 33657023, 33657109, 33657349, 35695769, 18108370, 18108374, 18108376, 35698423, 35695855, 264555, 52644332, 56182323, 60170394, 83373044, 56528488
2391	80036194 (4781, 4782)			UNCLASSIFIED	263978
2392	94245016 (4783, 4784)	Novel Protein sim. GBank gll4240169[dbj]BAA74863.11 - (AB020647) KIAA0840 protein [Homo sapiens]	Contains protein domain (PF00560) - Leucine Rich Repeat	nuclease	35696286, 35696052, 264508, 264905, 264509, 264908, 264907, 264908, 264909, 264510, 264511, 264512, 264910, 265009, 264591, 264758, 264600, 264604, 264762, 264448, 264764, 264369, 264768, 264768, 264769, 264689, 35695917, 264629, 18108374, 263978, 35696423, 35695855, 264631, 264634, 264635, 264638, 264637, 264638, 60170394, 264639, 264555, 264488

2393	95302833 (4785, 4786)	Novel Protein sim. GBank gll4506867[et]NP_000993.1[pr]PLP - ribosomal protein, large, P0	Contains protein domain (PF00468) - ribosomal prot Ribosomal protein L10	18108392, 60424178, 264489, 18108394, 18108397, 22278995, 58994075, 35896286, 22278996, 22278997, 22278999, 264093, 60432049, 264259, 29331822, 29147820, 20281099, 29331824, 29331825, 68714117, 60432289, 29331826, 29331827, 29331828, 35696052, 29146499, 264508, 264509, 264805, 264907, 264908, 68712502, 52844045, 264828, 264909, 56182435, 264112, 264113, 264510, 265008, 264511, 265007, 265008, 265009, 264910, 264591, 264593, 60433356, 264595, 60433438, 52846317, 33109954, 21906754, 55811388, 265010, 265011, 265017, 265018, 265019, 264681, 264762, 18108351, 264763, 264682, 264764, 264683, 264389, 264288, 18108354, 264766, 264688, 264687, 264688, 264689, 18108359, 21906765, 21906766, 21906767, 21906768, 21906769, 55811957, 29148629, 29148784, 35895917, 265021, 265022, 33657023, 264692, 264693, 18108384, 33657109, 18108368, 27486281, 27486282, 33657349, 35895763, 18108370, 263972, 264629, 18108374, 263977, 18108376, 263978, 55810764, 35696423, 35895855, 264634, 60431850, 264555, 264637, 264557, 263981, 264558, 18108381, 60170394, 35696286, 22278997, 22278998, 56182181, 35696052, 265006, 264592, 55811386, 265010, 265011, 265017, 265019, 264448, 264683, 264288, 21906765, 21906766, 21906769, 55811957, 35695917, 33657023, 65274620, 33657182, 33657349, 35895763, 18108374, 18108376, 55810764, 55811376, 35696423, 60170394, 18108385, 264564, 264566, 264567
2394	94323268 (4787, 4788)	Novel Protein sim. GBank gll4159888 (AC004908) - zinc finger protein from gene of uncertain exon structure; similar to Q99876 (PID:g3025333) [Homo sapiens]	Contains protein domain (PF00086) - Zinc finger, C2H2 type	dna_rna_bind
2395	95287212 (4789, 4790)	Novel Protein sim. GBank gll571275[et]AAD47638.1[AF16079] - (AF160798) calcium transporter CaT1 [Rattus norvegicus]		dna_rna_bind

2398	85086700 (4781, 4792)	Novel Protein sim. GBank gl 106322 pir 834087 - hypothetical protein (L1H 3' region) - human	Contains protein domain (PF00560) - Leucine Rich Repeat	nuclease	52646385, 18108397, 56182575, 35698286, 22278997, 22278999, 60432049, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 264108, 264907, 29331830, 66712502, 264110, 60170831, 264591, 33657402, 60433438, 55812038, 33109954, 21908754, 33657084, 87168474, 265017, 265018, 265019, 264760, 264448, 264288, 264768, 52844229, 21908768, 21908767, 265020, 265021, 60170815, 264692, 33657023, 65274620, 52645129, 33657182, 27488262, 27488264, 27488265, 264629, 18108374, 35698423, 35698855, 264631, 264596, 52844332, 264558, 83373044, 18108388, 87168518, 22279002, 264482
2397	87280854 (4793, 4794)				52644507, 52645156, 56182575, 264259, 29147820, 264905, 264907, 264908, 264909, 264910, 264758, 52644298, 264603, 264604, 264762, 264681, 264764, 18108337, 264789, 21908768, 264693, 264628, 264635, 264639, 264638, 264584
2398	88047689 (4795, 4796)	Novel Protein sim. GBank gl 3258609 (AC005178) - H53_GS1 [Homo sapiens]		UNCLASSIFIED	
2399	87738965 (4797, 4798)	Novel Protein sim. GBank gl 785117 (L41834) - nuclear protein [Ensis minor]		UNCLASSIFIED	56994075, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 66714117, 60432289, 264906, 29331830, 56182435, 264112, 264910, 33109954, 21908754, 87168474, 264600, 265017, 265018, 265019, 264764, 264765, 21908765, 21908768, 21908767, 21908769, 35695917, 265020, 265022, 60170815, 33657023, 18108370, 18108374, 264556, 60170394, 264558, 87168518, 22279000, 22279002, 264564, 264566, 264487
2400	91214118 (4799, 4800)	Novel Protein sim. GBank gl 2352822 gb AAB69285.1 - (AF008945) glucose-6-phosphatase [Haplochromis nubilus]		phosphatase	21908768, 52646842, 56994075, 33657182, 27486262, 52644296, 265017

2401	91214118 (4801, 4802)	Novel Protein sim. GBank gij2352822[gbjAAB9285.1] - (AF008945) glucose-6-phosphatase [Haplochromis nubilus]		phosphatase	52644507, 52645156, 52644229, 264688, 21908764, 21908765, 52646385, 52646842, 21908768, 21908767, 21908768, 22278995, 35695917, 56984075, 35696286, 22278998, 22278997, 265020, 22278998, 22278998, 264259, 33657023, 52645080, 264693, 29331824, 33657109, 52645129, 29331826, 33657182, 29331827, 35696052, 27486261, 27486262, 33656970, 33657349, 27486265, 35695763, 264108, 264905, 35696423, 35695855, 265006, 265007, 265008, 265009, 264637, 52644332, 55812038, 52646317, 18108385, 52644298, 87168474, 265010, 87168559, 80432113, 265017, 265018, 265019, 264563, 264288, 264907, 264908, 264909, 264566
2402	91221408 (4803, 4804)	Novel Protein sim. GBank gij4689258[gbjAAD27832.1]AF12185 - (AF121859) sorting nexin 9 [Homo sapiens]			
2403	94135432 (4805, 4806)	Novel Protein sim. GBank gij4929575[gbjAAD34048.1]AF15181 - (AF151811) CGI-53 protein [Homo sapiens]	Contains protein domain (PF000082) - C-type lysozyme/alpha-lactalbumin family		22278999, 35696052, 265018, 264688, 264693, 83373044, 264567
2404	95312605 (4807, 4808)	Novel Protein sim. GBank gij2313786 (AF016685) - similar to short chain-type dehydrogenases [Caenorhabditis elegans]	Contains protein domain (PF00108) - short chain dehydrogenase	dehydrogenase	35696286, 29331826, 35696052, 265008, 265018, 21908769, 264564
2405	94311851 (4809, 4810)	Novel Protein sim. GBank gij464178[gbjBAA03581] - (D14853) polyprotein [Hepatitis C virus]		UNCLASSIFIED	35696286, 29331822, 265007, 21806754, 265017, 265018, 265019, 264763, 264369, 21908765, 35695917, 265020, 265021, 52644150, 264693, 35695855, 264632, 52644332, 22279002
2406	88094501 (4811, 4812)	Novel Protein sim. GBank gij2773363 (AF041382) - microtubule binding protein D-CLIP-190 [Drosophila melanogaster]	Contains protein domain (PF01302) - CAP-Gly domain	struct	52646842, 22278994, 22278996, 35696286, 22278997, 22278998, 22278999, 264092, 264093, 60432049, 264259, 29331822, 29331824, 29331825, 29331827, 35696052, 29331828, 264102, 264106, 264908, 52644045, 265007, 265008, 265009, 264910, 264592, 60433356, 60433438, 33109954, 265010, 265011, 265018, 265019, 264389, 264885, 264686, 21908768, 21908769, 52644150, 264693, 52645129, 264628, 35696423, 264632, 56182323, 264639, 22279000, 22279002, 264563, 264685, 264686
2407	79465005 (4813, 4814)			UNCLASSIFIED	
2408	87391503 (4815, 4816)	Novel Protein sim. GBank gij423442[prijS33513 - gene Fil protein - mouse]		UNCLASSIFIED	264910, 265010, 264448, 264557

2409	94741770 (4817, 4818)	Novel Protein sim. GBank gl 1176601 sp P45966 YN26_CAEEL - HYPOTHETICAL 20.8 KD PROTEIN T09A5.6 IN CHROMOSOME III		UNCLASSIFIED	22278995, 22278996, 22278997, 264097, 29331822, 29331824, 29331827, 29148498, 52644045, 60433438, 33657084, 87168474, 264760, 21908787, 29148627, 29148629, 52644150, 33657023, 263967, 20281069, 18108374, 20281071, 58182323, 83373044, 18108385, 87168518 22278998, 264259
2410	87604860 (4819, 4820)	Novel Protein sim. GBank gl 4966262 gb AAC48052.2 - (U64849) Contains similarity to P1am domain: PF00648 (F- box), Score=28.7, E-value=4.3e-05, N=1 [Caenorhabditis elegans]	Contains protein domain (PF00080) - Copper/zinc superoxide dismutase (SODC)		
2411	87534633 (4821, 4822)	Novel Protein sim. GBank gl 3114713 (AF061348) - Edp1 protein [Mus musculus]		inf	29331824, 29331827, 29331828, 264764, 264369, 33657109, 58182323
2412	87778332 (4823, 4824)	Novel Protein sim. GBank gl 5410336 gb AAD43038.1 - (AF106685) myelin gene expression factor 2 [Homo sapiens]	Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	dna_rna_bind	22278998, 29331827, 264807, 265011, 265017, 265018, 265019, 18108351, 21906766, 265020, 33657109, 264559, 18108385
2413	94133820 (4825, 4826)	Novel Protein sim. GBank gl 5282705 emb CAB45778.1 - (AL080214) hypothetical protein [Homo sapiens]	Contains protein domain (PF00038) - Intermediate filament proteins	struct	264488, 264259, 29331826, 264508, 264905, 264509, 264806, 264907, 264908, 264510, 264511, 264512, 265008, 265009, 264910, 265011, 284682, 264784, 264768, 264686, 264788, 264689, 265021, 33657023, 18108370, 264628, 35695855, 264632, 264634, 264635, 264636, 83373044, 264563, 264584, 264585, 264586, 264587, 264488
2414	94312590 (4827, 4828)	Novel Protein sim. GBank gl 1082340 pir J52863 - DNA- binding protein R kappa B - human		ubiquitin	52645156, 52646842, 52646365, 18108398, 56182375, 22278994, 22278995, 58994075, 22278996, 35696286, 22278997, 22278998, 22278999, 264259, 52645080, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 35696052, 33656970, 29331830, 264909, 60433356, 33657402, 264594, 52646317, 21908754, 33657084, 265010, 87168559, 265017, 265018, 265019, 264369, 264684, 264687, 264688, 56181582, 21908764, 264689, 21908765, 21908766, 21908767, 29148627, 21908769, 265020, 265021, 60170815, 33657023, 264693, 52645129, 33657109, 33657182, 27486261, 27486262, 27486265, 33657349, 18108370, 60431528, 264629, 18108374, 18108376, 55810764, 264636, 52644332, 264638, 264558, 58182323, 83373044, 18108385, 87168518, 22279002

2415	88089002 (4829, 4830)	Novel Protein sim. GBank gij423915 pir A45438 - myosin I heavy chain - rat	Contains protein domain (PF00063) - struct	284259, 284808, 60433356, 33657402, 21908754, 285018, 284687, 284689, 21908769, 55811957, 285021, 284690, 284691, 33657023, 284693, 35698423, 56182323, 56526486, 284638
2416	94118356 (4831, 4832)	Novel Protein sim. GBank gij3025445 (AC004528) - R32184_1 [Homo sapiens]		
2417	87733334 (4833, 4834)	Novel Protein sim. GBank gij1084944 pir S54495 - hypothetical protein YPR021c - yeast (Saccharomyces cerevisiae)	Contains protein domain (PF00153) - Mitochondrial carrier proteins	284094, 29331822, 29331824, 29331827, 284369
2418	94234349 (4835, 4836)	Novel Protein sim. GBank gij1176572 sp P45895 YNA4 CAEEL - HYPOTHETICAL 91.0 KD PROTEIN PAR2.4 IN CHROMOSOME III	Contains protein domain (PF00411) - UNCLASSIFIED Ribosomal protein S11	56994075, 264091, 264259, 29331824, 29331825, 60432289, 29331828, 284805, 284907, 284511, 265009, 60432229, 21908754, 87188559, 265019, 284682, 21908768, 21906769, 265020, 265021, 33657023, 65274620, 18108370, 55811576, 284634, 60170394, 18108385, 22279000, 22279002, 284588
2419	82374248 (4837, 4838)	Novel Protein sim. GBank gij284006 pir S18732 - autoantigen, 64K - human	struct	284569, 264762, 264448, 284691, 264631, 264634, 264555, 264558, 284638, 284558
2420	94844244 (4839, 4840)	Novel Protein sim. GBank gij107621 pir S50755 - hypothetical protein VSP-3 - Chlamydomonas reinhardtii	UNCLASSIFIED	29331824, 29331825, 29331828, 60432229, 33109954, 85658542, 87168474, 285018, 284288, 265020, 284584
2421	87805345 (4841, 4842)	Novel Protein sim. GBank gij2224567 dbj BAA20772 (AB002311) KIAA0313 [Homo sapiens]	UNCLASSIFIED	264809, 284768, 284638
2422	88084714 (4843, 4844)	Novel Protein sim. GBank gij4505153 ref NP_002392.1 pMEKK - MAP/ERK kinase kinase 3	transport	18108392, 18108394, 18108398, 284806, 265008, 265010, 18108351, 18108374, 18108385
2423	88058380 (4845, 4846)	Novel Protein sim. GBank gij4505153 ref NP_002392.1 pMEKK - MAP/ERK kinase kinase 3	kinase	264259, 60432049, 29331822, 29331826, 60432289, 29331828, 265008, 265009, 60433356, 21908754, 265017, 265018, 285019, 21908766, 21906768, 21906769, 265020, 265021, 20281149, 263971, 60432113
2424	94854047 (4847, 4848)	Novel Protein sim. GBank gij2988398 (AC004391) - Unknown gene product [Homo sapiens]	UNCLASSIFIED	56182575, 35696286, 22278997, 60432049, 284259, 29331826, 29331828, 284805, 66712502, 29331830, 60433356, 265011, 285019, 284768, 21906768, 55811957, 284692, 33657023, 33657109, 55811576, 56182323, 83373044, 18108385, 18108388, 60432113, 22279000
2425	87415981 (4849, 4850)	Novel Protein sim. GBank gij2077832 dbj BAA19879 (D88558) Protein Kinase [Rattus norvegicus]	kinase	264634
2426	87813945 (4851, 4852)	Novel Protein sim. GBank gij2039388 gb AAAB53003.1 (U94619) circulating cathodic antigen [Schistosoma mansoni]	UNCLASSIFIED	222 8998, 22278998, 264259, 264102, 284512, 265008, 21908767, 18108370, 18108374, 263978

2427	87622693 (4853, 4854)	Novel Protein sim. GBank gij4680895jg/jAAD27737.1jAF13298 - (AF132982) CGI-28 protein [Homo sapiens]	Contains protein domain (PF00573) - Ribosomal protein L4/L1 family	ribosomalprot	264259, 20281099, 35696052, 265008, 264594, 265011, 264760, 18108351, 264882, 264883, 264369, 264684, 264688, 264687, 264689, 21908768, 264691, 264692, 18108374, 18108377, 264557, 264639, 18108385
2428	85732889 (4855, 4856)	Novel Protein sim. GBank gij1537070 (U63840) - nucleoporin p54 [Rattus norvegicus]			22278998, 22278999, 35696052, 21906754, 264288, 21908765, 21908768, 21908769, 35695917, 265020, 263972, 22278002
2429	87769276 (4857, 4858)	Novel Protein sim. GBank gij601931 (M94316) - neurofilament-H [Oryctolagus cuniculus]	Contains protein domain (PF00711) - Beta defensins	UNCLASSIFIED	22278999, 29331824, 264808, 264809, 284511, 265009, 21906754, 265017, 285018, 265019, 264448, 264683, 264288, 21908765, 21908768, 285021, 284693, 18108381
2430	86948827 (4859, 4860)	Novel Protein sim. GBank gij3860729jembjCAA14630] - (AJ235270) CELL DIVISION PROTEIN FTSJ (fts.j) [Rickettsia prowazekii]	Contains protein domain (PF01728) - FtsJ cell division protein	UNCLASSIFIED	264112, 264691
2431	87649884 (4861, 4862)	Novel Protein sim. GBank gij3876367jembjCAA93287] - (Z69360) Weak similarity to Elmeria thrombospondin (PIR Acc. No. A45517); cDNA EST EMBL:M89268 comes from this gene; cDNA EST yk295b9.5 comes from this gene [Caenorhabditis elegans]		protease	28331826, 29331827, 35696052, 29146489, 264905, 264906, 264681, 264288, 264689, 21908765, 264692, 35696423 264634, 264558
2432	80083033 (4863, 4864)	Novel Protein sim. GBank gij2224593jdbijBAA20784] - (AB002324) KIAA0326 [Homo sapiens]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	transcriptfactor	284569, 284905, 265018, 264762, 264683, 284691, 284558, 264557, 264639, 284558 264563
2433	80055082 (4865, 4866)	Novel Protein sim. GBank gij1263289 (U47856) - fibronin-4 [Araneus diadematus]		UNCLASSIFIED	284555
2434	19520148 (4867, 4868)	Novel Protein sim. GBank gij3641352 (AF091234) - putative transcription factor [Mus musculus]		UNCLASSIFIED	265008, 264758, 265010, 264689, 27486261, 283972, 18108374, 18108381
2435	20759044 (4869, 4870)	Novel Protein sim. GBank gij3860014 (AF091088) - unknown [Homo sapiens]	Contains protein domain (PF01256) - Uncharacterized protein family UPF0031	UNCLASSIFIED	29331828, 265007, 265009, 265017, 264750, 264685, 264693, 264565
2436	88044008 (4871, 4872)	Novel Protein sim. GBank gij3860014 (AF091088) - unknown [Homo sapiens]		UNCLASSIFIED	29331826, 264508, 264905, 264509, 264906, 264907, 264908, 264909, 264511, 265008, 264512, 264910, 265009, 264591, 33857402, 21908754, 265011, 264760, 264784, 264685, 264688, 264768, 35695917, 33857023, 284693, 264631, 264632, 58182323, 264558, 83373044, 264563, 264564, 264565, 264568, 264567

2439	94850850 (4877, 4878)	Novel Protein sim. GBank gi 4263519 gb AAD15345 - (AC004044) small nuclear riboprotein Sm-D1 (Arabidopsis thaliana)	Contains protein domain (PF01423) - Sm protein	UNCLASSIFIED	60424179, 18108397, 56182575, 56181886, 56994075, 22278998, 35696286, 22278997, 22278999, 264259, 52645080, 29331822, 56182181, 29331824, 80424289, 68714117, 29331825, 60432289, 29331826, 29331827, 29331828, 35698052, 29146498, 264509, 264905, 264906, 52644045, 80431735, 33109954, 21908754, 33657084, 55811388, 52844298, 87188474, 265017, 265018, 265019, 18108351, 264448, 264288, 264788, 52644229, 56181562, 21908764, 21908765, 21908768, 21908767, 21908768, 21908769, 35695917, 33657023, 33657108, 33657182, 27486282, 27486284, 33657349, 27486285, 35695763, 18108370, 60431528, 263977, 55810764, 35698423, 65274791, 35695655, 60431850, 56182323, 60432113, 22279000, 22279002, 264567
2440	87641733 (4879, 4880)			UNCLASSIFIED	29331822, 29331824, 29331825, 29331827, 35696052, 264508, 264907, 264510, 265018, 265019, 264448, 264369, 265020, 265021, 56182323, 264639, 22279002
2441	87823914 (4881, 4882)	Novel Protein sim. GBank gi 3024889 sp P56524 Y288_HUMAN - HYPOTHETICAL PROTEIN KIAA0288 (HA8116)		UNCLASSIFIED	264488, 264629, 18108374, 264584
2442	87273580 (4883, 4884)	Novel Protein sim. GBank gi 4506013 ref NP_002703.1 pPPP1 - protein phosphatase 1, regulatory subunit 7	Contains protein domain (PF00360) - Leucine Rich Repeat	UNCLASSIFIED	22278996, 22278999, 29331822, 264768, 264693
2443	84305949 (4885, 4886)	Novel Protein sim. GBank gi 1170658 sp Q02875 KID1_RAT - RENAL TRANSCRIPTION FACTOR KID-1 (TRANSCRIPTION FACTOR 17)	Contains protein domain (PF01352) - KRAB box	transcription factor	264908
2444	88086345 (4887, 4888)	Novel Protein sim. GBank gi 4758824 ref NP_004280.1 pNRF3 - nuclear factor (erythroid-derived 2)-like 3	Contains protein domain (PF00170) - bZIP transcription factor	transcription factor	264259, 18108382, 18108383, 18108385, 22278000
2445	87338638 (4889, 4890)	Novel Protein sim. GBank gi 2135950 pir IS58222 - PQ-rich protein - human			264259, 35698052, 264369, 18108381
2446	88059293 (4891, 4892)	Novel Protein sim. GBank gi 4753887 emb CAA05409.2 - (AJ002424) p85 protein [Rattus norvegicus]	Contains protein domain (PF00085) - WAP-type (Whey Acidic Protein) 'four-disulfide core'	protease inhib	265011, 264689, 33657023, 263981, 18108385
2447	94845149 (4893, 4894)	Novel Protein sim. GBank gi 4885613 ref NP_005409.1 pST5 - suppression of tumorigenicity 5		cadherin	56182575, 264259, 29331824, 29331825, 29331827, 60433356, 60433438, 264758, 265018, 264692, 65274620, 60431528, 65274791, 56182323

2448	87749680 (4895, 4896)			UNCLASSIFIED	22278996, 22278997, 22278999, 28331828, 35698052, 264107, 264110, 87168474, 87168559, 18108351, 21906767, 21906769, 21486282, 263978
2449	87869075 (4897, 4898)	Novel Protein sim. GBank gl 728837 sp P39194 ALU7_SQ WARNING ENTRY IIII	HUMAN - IIII ALU SUBFAMILY	cadherin	264259, 284828, 265007, 264595, 265021, 56526488
2450	86597784 (4899, 4900)			UNCLASSIFIED	264906
2451	91014503 (4901, 4902)	Novel Protein sim. GBank gl 1710021 sp P35290 RB24_PROTEIN RAB-24 (RAB-16)	Contains protein domain (PF00071) - Ras family	glycoprotein	264093, 29331822, 29331824, 29331825, 66714117, 29331826, 29331828, 35698052, 264907, 66712502, 29331830, 264910, 265009, 284758, 265017, 265018, 264782, 264448, 284288, 21906767, 265021, 33657023, 264693, 33657109, 263969, 83373044, 18108385
2452	91230509 (4903, 4904)	Novel Protein sim. GBank gl 1504034 dbj BAA13216 - (D86980) KIAA0227 [Homo sapiens]		isomerase	264102, 284112, 284688, 263972, 18108374, 83373044, 264583
2453	84201088 (4905, 4906)	Novel Protein sim. GBank gl 2880078 (AC004142) - similar to murine leucine-rich repeat protein; possible role in neural development by protein-protein interactions; 93% similarity to D49802 (PID:gl 369908) [Homo sapiens]	Contains protein domain (PF00560) - Leucine Rich Repeat	nglrecep	264509, 264512, 18108385
2454	95310691 (4907, 4908)	Novel Protein sim. GBank gl 1076802 pir J548915 - extensin like protein - maize	Contains protein domain (PF00170) - bZIP transcription factor	UNCLASSIFIED	263994, 66714117, 29331827, 264508, 264509, 284905, 264908, 264907, 284908, 264909, 264510, 264511, 264512, 265009, 264910, 264591, 264758, 264759, 265010, 265011, 284603, 264604, 264760, 284781, 264782, 18108351, 264764, 264765, 264768, 264886, 264788, 264769, 264534, 284691, 264692, 33657023, 284693, 33657109, 284628, 263978, 35698055, 284634, 264635, 264637, 264638, 264639, 83373044, 18108385, 284563, 264564, 264488
2455	95288301 (4909, 4910)	Novel Protein sim. GBank gl 543817 sp P35585 AP47_MOUSE - CLATHRIN COAT ASSEMBLY PROTEIN AP47 (CLATHRIN COAT ASSOCIATED PROTEIN AP47) (GOLGI ADAPTOR AP-1 47 KD PROTEIN) (HA1 47 KD SUBUNIT) (CLATHRIN ASSEMBLY PROTEIN ASSEMBLY PROTEIN COMPLEX 1 MEDIUM CHAIN)	Contains protein domain (PF00928) - Adaptor complexes medium subunit family	glycoprotein	264488, 22278996, 264259, 35698052, 264905, 264908, 264907, 264908, 264909, 264510, 264512, 265008, 265009, 264910, 264591, 264592, 264593, 264594, 264758, 265019, 264760, 264681, 18108351, 264683, 264784, 18108354, 264766, 264768, 264789, 264689, 21906768, 21906767, 21906769, 29148629, 35698057, 265020, 265022, 33657023, 33657109, 18108370, 264628, 264629, 264631, 264632, 264635, 58182323, 60170394, 18108385, 264563, 284584, 264566, 264567
2458	88166700 (4911, 4912)	Novel Protein sim. GBank gl 2388630 (AC003079) - Ankyrin-like; 54% similar to 2022340A (NID:gl 1092123) in exons spanning 43974 to 11551 of clone. [Homo sapiens]	Contains protein domain (PF00023) - Ank repeat	kinase	264693

2457	84118375 (4913, 4914)	Novel Protein sim. GBank gij3025447 (AC004528) - R32184_3 [Homo sapiens]		UNCLASSIFIED	55181686, 264805, 264907, 264511, 264596, 55811386, 264882, 264684, 264685, 264887, 264691, 33657023, 264693, 35695855, 264636, 264555, 56182323, 264558, 56528486, 264563
2458	85875304 (4915, 4916)	Novel Protein sim. GBank gij2384942 (AF022985) - Similar to collagen [Caenorhabditis elegans]		UNCLASSIFIED	264691, 264693, 264634, 264559
2459	87551913 (4917, 4918)	Novel Protein sim. GBank gij5441842 [gb]AAD43187.1 [AC004997] supported by mouse EST AA538043 (NID:g2284036) [Homo sapiens]	Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger)	transport	27486265
2460	84315289 (4919, 4920)	Novel Protein sim. GBank gij4928701 [gb]AAD34111.1 [AF15187 - (AF151874) CGI-116 protein [Homo sapiens]		kinase	65274572, 35696286, 22278996, 22278997, 80432049, 29331822, 29331824, 29331825, 66714117, 60432289, 29331826, 29331827, 29331828, 33658970, 29146499, 264102, 264109, 60433438, 265017, 265018, 265019, 264288, 21808765, 21806766, 21808769, 35695817, 265020, 264691, 33657023, 27486261, 18108374, 35695855, 87188518, 60432113
2461	87645147 (4921, 4922)	Novel Protein sim. GBank gij4426862 [gb]AAD20633 - (AF128062) Arf-like 2 binding protein BART1 [Homo sapiens]		UNCLASSIFIED	264259, 29331828, 264910, 18108351, 18108370, 18108374
2462	86988002 (4923, 4924)	Novel Protein sim. GBank gij5420387 [emb]CAB46679.1 - (AJ243459) prolephosphoglycan [Leishmania major]			264909, 264758, 264684, 18108374, 264637, 18108385
2463	84388543 (4925, 4926)	Novel Protein sim. GBank gij5052516 [gb]AAD38568.1 [AF14561 - (AF145813) BcDNA, GH03108 [Drosophila melanogaster]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	UNCLASSIFIED	264681, 264566
2464	91218957 (4927, 4928)	Novel Protein sim. GBank gij5410300 [gb]AAD43021.1 - (AF100757) COP9 complex subunit 4 [Homo sapiens]	Contains protein domain (PF01399) - PCI domain	protease	264489, 52646842, 22278995, 35696286, 22278996, 22278997, 22278998, 264259, 29331822, 29331824, 66714117, 29331825, 29331826, 29331827, 35698052, 29331828, 264509, 52644045, 264510, 264511, 264512, 265008, 60170831, 264593, 52646317, 33109954, 33657084, 265017, 265018, 265019, 264782, 264448, 264764, 264288, 264766, 21906765, 21906766, 21908767, 21908768, 21906769, 265021, 33657023, 33657109, 18108370, 18108381, 60170394, 18108385, 22279002, 264486

2465	95357483 (4929, 4930)	Novel Protein sim. GBank gij4508401[re]NP_002871.1[prAF1 - v-rafl murine leukemia viral oncogene homolog 1	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	oncogene	18108392, 52644507, 52845156, 52846365, 22278994, 22278995, 35696286, 22278998, 22278998, 284259, 29331822, 29331824, 29331825, 60424269, 60432289, 29331827, 35896052, 29331828, 284907, 29331830, 52844045, 284909, 56182435, 264511, 265007, 265008, 265009, 264910, 33657402, 60433438, 55812038, 21908754, 33109954, 285010, 265011, 87168559, 284600, 285017, 265018, 285019, 18108351, 284369, 284288, 284685, 284767, 21908765, 21908767, 21908768, 55811957, 35695917, 265020, 265021, 60170815, 52844150, 33657023, 52845129, 33657109, 27486281, 27486284, 35695763, 284628, 263972, 18108374, 35695855, 284638, 284637, 60170394, 56528488, 87168518, 60432113, 284563, 284564, 284566, 284487 264369
2466	85681388 (4931, 4932)	Novel Protein sim. GBank gij4321819[gb]AAD15788.1] - (AF051098) seven transmembrane domain orphan receptor [Mus musculus]			
2467	88059485 (4933, 4934)	Novel Protein sim. GBank gij3513300 (AC005595) - F16801.1, partial CDS [Homo sapiens]		UNCLASSIFIED	56994075, 264908, 21908768, 33657023
2468	87614698 (4935, 4936)	Novel Protein sim. GBank gij2143455[pil]i58106 - gene DMR-N9 protein - mouse (fragment)	Contains protein domain (PF00400) - WD domain, G-beta repeat	kinase	29331824, 52844045, 265008, 284910, 265019, 21908765, 21908768, 265021
2469	86294397 (4937, 4938)	Novel Protein sim. GBank gij5420389[emb]CAB46680.1] - (AJ243460) proteophosphoglycan [Leishmania major]		UNCLASSIFIED	284288, 284628
2470	80223831 (4939, 4940)	Novel Protein sim. GBank gij5419882[emb]CAB46424.1] - (AL098749) DKFZp434G153 [Homo sapiens]		UNCLASSIFIED	264592, 264692, 264555, 264556, 264557, 264558, 264559, 18108385, 284482
2471	91013681 (4941, 4942)	Novel Protein sim. GBank gij5419882[emb]CAB46424.1] - (AL098749) DKFZp434G153 [Homo sapiens]		UNCLASSIFIED	85274572, 35896286, 29331827, 285007, 264592, 33109954, 265018, 265019, 284288, 55811957, 265020, 264693, 55811576, 56182323
2472	95060811 (4943, 4944)	Novel Protein sim. GBank gij4929747[gb]AAD34134.1[AF15189 - (AF151897) CGI-139 protein [Homo sapiens]		UNCLASSIFIED	264092, 60432049, 29331825, 60433356, 265010, 265011, 18108351, 264784, 284288, 264692, 65274620, 18108370, 18108372, 18108374, 284634, 18108385
2473	95421508 (4945, 4946)	Novel Protein sim. GBank gij4539009[emb]CAB39630.1] - (AL049481) putative protein [Arabidopsis thaliana]			80424178, 65274572, 22278999, 60424269, 29331828, 265008, 60433356, 60433438, 265010, 18108351, 284448, 264288, 284687, 264689, 265021, 264692, 65274620, 60431528, 65274791, 264556, 56182323, 60432113
2474	94315616 (4947, 4948)	Novel Protein sim. GBank gij3252827 (AC004382) - Unknown gene product [Homo sapiens]			65274572, 56994075, 264259, 29331826, 60170831, 285017, 265018, 265019, 284683, 284369, 265020, 264693, 264563, 264564

2475	94321693 (4949, 4950)	Novel Protein sim. GBank gi 1216486 (U48852) - HT protein [Cricetus gresus]	Contains protein domain EGF-like domain	igf	264259, 29331822, 265006, 265007, 265010, 265011, 264448, 264288, 264369, 264885, 264886, 18108357, 264768, 18108362, 264693, 18108370, 18108374, 18108379, 35698423, 83373044, 18108383, 18108385, 264564, 264585, 264567
2476	94315818 (4951, 4952)	Novel Protein sim. GBank gi 3252827 (AC004382) - Unknown gene product [Homo sapiens]		UNCLASSIFIED	264259, 60424269, 66714117, 264905, 265006, 264511, 265008, 265009, 264758, 265010, 265011, 18108351, 264681, 264369, 264288, 264689, 21906767, 265020, 18108374, 264639, 18108382, 83373044, 18108385, 87168518
2477	20718974 (4953, 4954)			UNCLASSIFIED	263978
2478	17659165 (4955, 4956)			UNCLASSIFIED	265017
2479	94314569 (4957, 4958)	Novel Protein sim. GBank gi 1644232[dbj][BAA11082] - (D67086) N-WASP [Bos taurus]		im7	56994075, 22278999, 21906754, 264682, 21908765
2480	95293605 (4959, 4960)			UNCLASSIFIED	264905, 264907, 264765
2481	94718481 (4961, 4962)	Novel Protein sim. GBank gi 5689469[dbj][BAA03018.1] - (AB028989) KIAA1066 protein [Homo sapiens]		collagen	65274572, 56182575, 22278997, 284094, 264259, 29331822, 29331824, 66714117, 29331827, 35698052, 264508, 284905, 264906, 264907, 264908, 52644045, 264909, 56182435, 265008, 264910, 33657402, 55812038, 264758, 265010, 265011, 265017, 265018, 264760, 264762, 18108351, 284784, 284288, 264768, 264688, 264768, 21906768, 55811957, 265020, 264691, 264692, 264693, 284628, 55811576, 264630, 264634, 264635, 264636, 264637, 264556, 264558, 56182323, 83373044, 60432113, 22278002
2482	87383165 (4963, 4964)	Novel Protein sim. GBank gi 321249[pir][S28407] - guanine nucleotide-exchange activator CDC25 homolog - mouse		UNCLASSIFIED	29331822, 29331824, 29331825, 29331827, 264508, 264905, 264509, 264906, 264907, 264908, 264511, 264591, 284768, 264693, 264631, 264632, 264636, 264638, 264639, 284583
2483	87731583 (4965, 4966)			UNCLASSIFIED	264488, 22278995, 264093, 264095, 60432049, 60433356, 60433438, 264448, 264288, 263967, 18108370, 18108385, 18108388, 264482
2484	94187774 (4967, 4968)	Novel Protein sim. GBank gi 728831[sp][P39188]ALU1_HUMAN - IIII ALU SUBFAMILY J WARNING ENTRY IIII		kinase	264563
2485	87766556 (4969, 4970)	Novel Protein sim. GBank gi 185397 (U25281) - SH3 domain binding protein [Rattus norvegicus]		UNCLASSIFIED	22278995, 22278996, 22278997, 22278999, 264259, 60432049, 29331824, 60432289, 29331827, 265007, 264910, 264593, 264600, 264603, 264604, 265018, 264448, 264288, 264685, 264686, 264769, 264689, 35685917, 265022, 264692, 264693, 56182323

2486	87748978 (4971, 4972)	Novel Protein sim. GBank gi 2662167 dbj BAA33715 - (AB007803) KIAA0443 [Homo sapiens]			265017, 264555
2487	95343105 (4973, 4974)	Novel Protein sim. GBank gi 464559 sp P35287 RB14_RAT RAS-RELATED PROTEIN RAB-14	Contains protein domain (PF00071) - Ras family	glycoprotein	22278996, 35696286, 22278997, 22278998, 22278999, 264092, 264259, 29331822, 35696052, 264108, 264905, 264907, 29331830, 264908, 265008, 264511, 265008, 265009, 60433438, 21906754, 33109954, 87188559, 265018, 264681, 264288, 264687, 21908765, 21908766, 21908767, 21908768, 21908769, 35695917, 265021, 265022, 264534, 33857023, 264692, 33857109, 263972, 18108377, 35698423, 35698655, 60170394, 18108385, 56528488, 22278000, 22279002, 264583, 264482, 264585, 20281189, 18108391
2488	87652451 (4975, 4976)			UNCLASSIFIED	264910, 264448, 264288, 264584, 264691, 264634
2489	82990585 (4977, 4978)	Novel Protein sim. GBank gi 4886439 emb CAB43355.1 - (AL050253) hypothetical protein [Homo sapiens]	Contains protein domain (PF00084) - Sushi domain (SCR repeat)	complementrecept	264686, 264693, 55811576, 22279002
2490	88068609 (4978, 4980)	Novel Protein sim. GBank gi 2588624 (AC003083) - Rep2 Interacting protein-like: similar to U73941 (PID:g1816018) [Homo sapiens]		UNCLASSIFIED	264907, 265008, 22279002
2491	91242116 (4981, 4982)	Novel Protein sim. GBank gi 728832 sp P39189 ALU2_HUMAN - IIII ALU SUBFAMILY SB WARNING ENTRY IIII		Im7	264259, 29331826, 265008, 264782, 18108370, 18108376, 18108379
2492	95308202 (4983, 4984)	Novel Protein sim. GBank gi 3355303 (AF001549) - Unknown gene product [Homo sapiens]		transcriptionfactor	264488, 22278998, 22278999, 29331828, 264591, 33109954, 265017, 55811150, 21908764, 21908768, 264692, 60431528, 87188518, 60432113, 22279000

2483	95422415 (4985, 4986)	Novel Protein sim. GBank gi 4240307 dbj BAA74932.1 - (AB020716) KIAA0809 protein [Homo sapiens]	Contains protein domain (PF01424) - struct R3H domain	18108394, 264887, 65274572, 58182575, 22278995, 56994075, 60432049, 29331822, 29331824, 29331825, 29331826, 29331827, 29146498, 264508, 264905, 264509, 264908, 264907, 29331830, 264909, 264510, 265008, 264511, 265007, 264512, 265008, 265009, 264910, 21908754, 265011, 264600, 265017, 265018, 264604, 264805, 265019, 55811150, 264762, 18108351, 264681, 264448, 264683, 264389, 264288, 18108355, 18108357, 264687, 21908765, 21908766, 21908767, 21908768, 21908769, 265020, 264691, 264692, 33657023, 33657349, 18108370, 18108374, 18108376, 55810784, 18108379, 65274791, 264630, 264632, 264634, 264635, 264636, 264555, 264637, 264557, 264558, 264639, 264559, 83373044, 18108385, 87188518, 60432113, 22279000, 22279002, 264482, 264588, 264486
2484	30793118 (4987, 4988)			
2485	94234551 (4989, 4990)	Novel Protein sim. GBank gi 5420389 emb CAB46680.1 - (AJ243460) proteophosphoglycan [Leishmania major]	UNCLASSIFIED collagen	263994, 22278997, 35696032, 284509, 264905, 264908, 264907, 264908, 264909, 265006, 265009, 264595, 264604, 264448, 264682, 264764, 264288, 264685, 264788, 264789, 264689, 265020, 264692, 65274820, 264629, 55810784, 35696423, 55811578, 264636, 264637, 18108385, 22279000, 264584, 264587, 264486
2486	80018765 (4991, 4992)	Novel Protein sim. GBank gi 4808220 emb CAB42832.1 - (AL022315) dJ117715.1 (PUTATIVE novel protein) [Homo sapiens]	struct	29147620, 264805, 265008, 265007, 18108348, 18108362, 18108370, 18108374, 264555, 264556, 18108381, 18108383, 18108388
2487	81723554 (4993, 4994)		UNCLASSIFIED	52644507, 22278986, 22278989, 28331824, 29331828, 33657402, 21908754, 87168474, 265019, 264369, 264689, 21908785, 21908768, 21908787, 21908768, 265020, 33657023, 18108376, 18108387
2488	87724633 (4995, 4996)	Novel Protein sim. GBank gi 1200503 (U47924) - B [Homo sapiens]	UNCLASSIFIED	29331827, 264512, 264910, 264288, 18108374, 35695855
2489	94685125 (4997, 4998)	Novel Protein sim. GBank gi 3510234 (AC005581) - R31237_1, partial CDS [Homo sapiens]	Contains protein domain (PF00069) - kinase Eukaryotic protein kinase domain	264909, 55812038, 264631, 264637, 264558

2500	94649324 (4999, 5000)	Novel Protein sim. GBank gl 3881275 emb CAA21725 - (AL032655) predicted using GeneFinder; similar to Inositol monophosphatase family; cDNA EST YK255e11.5 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00459) - Inositol monophosphatase family	transport	52844507, 52645158, 22278995, 56994075, 35898288, 22278998, 264259, 52845080, 29331824, 29331825, 66714117, 60432289, 29331828, 29331827, 35696032, 29331828, 264508, 264509, 264510, 264512, 33657402, 60433438, 21906754, 52644298, 67168474, 87168559, 264603, 264681, 264448, 264683, 264288, 264389, 52644229, 264689, 21906765, 21906766, 21906767, 21906768, 21906769, 55811957, 35695917, 265020, 265021, 52644150, 33657023, 264693, 33657182, 35695763, 35696423, 35695855, 52644332, 83373044, 18108387, 87168518, 22279002
2501	94303896 (5001, 5002)	Novel Protein sim. GBank gl 4929815 gb AAD34068.1 AF15183 - (AF151831) CGI-73 protein [Homo sapiens]	Contains protein domain (PF00651) - BTB/POZ domain	dna_rna_bind	65274572, 58182575, 35696286, 22278996, 56994075, 22278997, 60432049, 264259, 29331822, 29331824, 29331826, 29331827, 35696052, 264905, 264906, 264907, 264908, 264909, 58182435, 264510, 264511, 265007, 264910, 264591, 60432229, 33657402, 60433356, 264595, 55812038, 264758, 264598, 87168474, 87168559, 264600, 264601, 264602, 265017, 264604, 265018, 264605, 265019, 18108351, 264448, 264369, 264288, 264768, 18108357, 21906765, 21906766, 21906767, 21906769, 29148629, 35695917, 264692, 33657023, 264629, 35698423, 55811576, 35695855, 264630, 264634, 264635, 264555, 264636, 264638, 264558, 60170394, 83373044, 18108385, 18108387, 87168518, 60432113, 22279002, 264586
2502	90993718 (5003, 5004)	Novel Protein sim. GBank gl 3041847 (AC004542) - OXYSTEROL-BINDING PROTEIN-like; similar to P22059 (PID:g129308) [Homo sapiens]	Contains protein domain (PF01237) - Oxysterol-binding protein	UNCLASSIFIED	65274572, 264907, 56182435, 265007, 264592, 264760, 18108351, 264448, 264369, 264288, 264684, 264686, 55811957, 265021, 264692, 33657109, 263973, 55811576, 264635, 264555, 264556, 264557, 264558, 56182323, 264558, 87168518, 264583, 264482
2503	87878345 (5005, 5006)	Novel Protein sim. GBank gl 2196874 emb CAA72638 - (Y11898) BRX protein [Mus musculus]			264905, 264907, 264512, 265008, 265011, 18108351, 264448, 264288, 29148627, 264693, 18108370, 18108374, 18108385

2504	87868708 (5007, 5008)	Novel Protein sim. GBank gij550420[embjCAA48220] - (x68101) lrg [Rattus norvegicus]			284488, 52844507, 52645156, 52646842, 22278994, 264259, 52645080, 29331822, 29331824, 29331825, 29331826, 29331827, 35698052, 264908, 264908, 52844045, 265009, 60433358, 33657402, 60433438, 264595, 33109954, 87169474, 285017, 265019, 264448, 264288, 264768, 52644229, 21908785, 21906766, 21908767, 21908768, 52644150, 264692, 27486261, 27486262, 27486264, 27486265, 35695763, 35698423, 35695855, 52644332, 56182323, 18108387, 87169519, 60432113, 22279002, 264584
2505	8760559 (5009, 5010)			UNCLASSIFIED	264605
2506	81232328 (5011, 5012)	Novel Protein sim. GBank gij2137562[pirj]49635 - mouse Dhml protein - mouse		nuclease	284488, 52844507, 52645156, 52646385, 65274572, 22278995, 56994075, 22278996, 22278997, 22278998, 22278999, 284259, 60432049, 29331822, 29331825, 29331826, 29331828, 264509, 56182435, 264112, 264593, 60433358, 55812036, 21906754, 265011, 265017, 265018, 265019, 264805, 264762, 18108351, 264448, 264288, 264768, 21908765, 21908766, 21908767, 21908768, 21908769, 35695917, 265020, 265021, 265022, 60170815, 33657023, 27486264, 18108378, 35695855, 264637, 83373044, 18108385, 87168518, 60432113, 22279000, 264583, 264482, 264585
2507	95316233 (5013, 5014)	Novel Protein sim. GBank gij5174489[refjNP_006035.1]pK1AA - histone deacetylase 6	Contains protein domain (PF00850) - Histone deacetylase family	histone	264488, 263994, 264592, 264595, 264369, 264686, 264768, 35695917, 35696423, 264583
2508	95315505 (5015, 5016)	Novel Protein sim. GBank gij4826433[embjCAB42889.1] - (AL031447) dJ126A5.2.1 (novel protein) (isoform 1) [Homo sapiens]		UNCLASSIFIED	22278995, 22278999, 60432049, 264259, 29331828, 265006, 265007, 60433438, 33657084, 265010, 265017, 265018, 265019, 18108351, 264448, 18108354, 264369, 18108359, 21908765, 21908769, 55811857, 265020, 265022, 27486261, 33657349, 18108377, 35695855, 60432113, 22279002, 264583, 264585
2509	87813741 (5017, 5018)	Novel Protein sim. GBank gij1263289 (U47856) - fibrin-4 [Araneus diadematus]		UNCLASSIFIED	265007, 265008, 18108357, 264556, 264587

2510	95421378 (5018, 5020)	Novel Protein sim. GBank gl[32933537]gb[AAAC25762.1] - (AF071059) zinc finger RNA binding protein [Mus musculus]			dna_ma_bind	65274572, 22278994, 22278996, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 29331825, 66714117, 60432289, 29331826, 33656970, 264908, 66712502, 265007, 264910, 60170831, 60432229, 60433356, 60433438, 21906754, 87188474, 265017, 265018, 264448, 264288, 21906787, 21906788, 21906769, 55811957, 35695917, 265020, 265022, 264691, 33857023, 264693, 65274820, 33657109, 33657182, 27486282, 33657349, 18108370, 35695855, 264555, 56182323, 83373044, 60432113, 22279002, 265017, 21906764, 265020, 264692
2511	87384281 (5021, 5022)	Novel Protein sim. GBank gl[4323152]gb[AAAD16228.1] - (AF098863) Ets-protein Spi-C [Mus musculus]				
2512	88084771 (5023, 5024)	Novel Protein sim. GBank gl[4502075]ref[NP_001135.1]pAMFR - autocrine motility factor receptor	Contains protein domain (PF000097) - Zinc finger, C3HC4 type (RING finger)		transport	22278999, 264259, 29331825, 29331826, 29146499, 264907, 264909, 265008, 264591, 60432229, 21906754, 264763, 264883, 264766, 18108357, 264889, 21906769, 264693, 18108370, 263972, 18108374, 264558, 22279000
2513	95357843 (5025, 5026)	Novel Protein sim. GBank gl[3004657 (AF017777)] - bobby sox [Drosophila melanogaster]			UNCLASSIFIED	60424179, 52645156, 18108394, 22278994, 35696286, 56994075, 22278996, 29331822, 29331824, 60424269, 29331825, 29331827, 33656970, 60431735, 33657084, 87188559, 265017, 264448, 264369, 56181562, 21906766, 21906767, 21906768, 21906769, 265020, 265021, 33657023, 18108386, 33657109, 27486281, 27486282, 33657349, 18108374, 55810764, 35696423, 56182323, 264558, 18108385, 264510
2514	88094578 (5027, 5028)	Novel Protein sim. GBank gl[2258437 (AF008197)] - syncollin [Rattus norvegicus]			UNCLASSIFIED	
2515	67994509 (5029, 5030)	Novel Protein sim. GBank gl[3757727]emb[CAA18783] - (AL022727) dJ8019.7 (olfactory receptor-like protein [hs6M1-3]) [Homo sapiens]	Contains protein domain (PF00001) - 7 transmembrane receptor (rhodopsin family)		Im7	
2516	87766908 (5031, 5032)				UNCLASSIFIED	264259, 28146498, 264905, 264288, 29148629, 35695917, 27486281, 264634
2517	87784968 (5033, 5034)	Novel Protein sim. GBank gl[4220527]emb[CAA23000] - (AL035356) putative protein [Arabidopsis thaliana]			UNCLASSIFIED	264091, 29331824, 29331825, 29331826, 29331828, 35696052, 264508, 264509, 264905, 264908, 264907, 264908, 264909, 264511, 264910, 33657402, 264757, 33109954, 265017, 265018, 264605, 264760, 264762, 264763, 264766, 264768, 264769, 33657109, 33657182, 264628, 55811578, 35696423, 264631, 264634, 264637, 264638, 264639, 87168518, 22279002, 264564

2518	94147410 (5035, 5038)	Novel Protein sim. GBank gi 4928591 gb AAD34056.1 AF151818 (AF151818) CGI-81 protein [Homo sapiens]	Contains protein domain (PF00018) - SH3 domain	UNCLASSIFIED	35686286, 264259, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 35686052, 29331828, 264907, 264909, 264511, 265007, 60432229, 60433356, 60433438, 55812038, 265010, 265017, 264448, 264288, 264689, 21906769, 21906769, 265022, 52644150, 264693, 18108370, 263972, 264555, 56182323, 83373044, 18108385, 60432113, 264088, 264259, 66714117, 29331826, 29331827, 29331828, 264907, 68712502, 265008, 265008, 264594, 265010, 265011, 265018, 264288, 21906769, 265020, 60431528, 55811576, 65274791, 264632, 264555, 264636, 22278002, 264564
2519	94328180 (5037, 5038)	Novel Protein sim. GBank gi 4263748 gb AAD15420 - (AC004883) similar to KIAA0766; similar to PID:g3682253 [Homo sapiens]		kinase	
2520	87413235 (5039, 5040)	Novel Protein sim. GBank gi 4826722 ref NP_005085.1 pFATP - fatty acid transport protein 4		transport	264259, 264908, 264910, 264682, 21906769, 265020, 264563
2521	85318244 (5041, 5042)	Novel Protein sim. GBank gi 5174488 ref NP_008035.1 pKIAA - histone deacetylase 6	Contains protein domain (PF00850) - Histone deacetylase family	histone	264488, 264489, 263994, 65274572, 22278995, 22278998, 264259, 29331822, 29331826, 264508, 264905, 264509, 264906, 264907, 66712502, 264511, 265008, 265007, 264591, 264592, 264593, 264594, 264595, 264596, 264681, 264448, 264763, 264682, 264764, 264684, 264369, 264288, 264685, 264686, 21906769, 55811957, 264682, 264693, 27488261, 18108370, 264628, 264628, 18108374, 55811576, 35696423, 35695855, 264632, 264558, 18108385, 65274727, 60432113, 264563, 264564, 264565, 264566, 264567
2522	87754052 (5043, 5044)	Novel Protein sim. GBank gi 4580011 gb AAD24201.1 U81002_ - (U81002) TRAF4 associated factor 1 [Homo sapiens]		transport	264489, 22278997, 20281171, 21906754, 35695917, 263987, 263976, 263981, 20281169
2523	85340467 (5045, 5046)	Novel Protein sim. GBank gi 1809327 (U76374) - skm- BOP2 [Mus musculus]	Contains protein domain (PF01753) - MYND finger		263969
2524	95340469 (5047, 5048)				56994075, 22278986, 35696286, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 29331826, 29331828, 29331830, 56182435, 264512, 265008, 60170831, 33657402, 265010, 87168559, 265019, 264288, 21906765, 21906769, 35695917, 265020, 265021, 265022, 52644150, 264691, 33657023, 33657109, 27488261, 35698423, 65274791, 264559, 83373044, 56528486, 87188518, 264567

2525	94126928 (5049, 5050)	Novel Protein sim. GBank gi 2073564 (U80223) - eukaryotic initiation factor eIF-2 alpha kinase; DGCNZ [Drosophila melanogaster]				264488, 22278997, 22278999, 60432048, 60432289, 29331828, 264905, 265008, 55812038, 21908754, 265019, 284369, 21908785, 21908788, 21908787, 21908769, 35695917, 265020, 265021, 33857109, 60431528, 83373044, 60432113, 22279000, 22279002, 284565
2526	95289404 (5051, 5052)	Novel Protein sim. GBank gi 4589828 (dbjBAA76836.1) - (AB023209) KIAA0992 protein [Homo sapiens]	Contains protein domain (PF00238) - Ribosomal protein L14	ribosomalprot		60424179, 284768, 264687, 264769, 264699, 65274572, 21908767, 58182575, 21908768, 21908769, 55811957, 22278994, 22278995, 35696286, 35695917, 22278996, 22278997, 265020, 22278998, 265021, 22278999, 265022, 264690, 264691, 60432049, 284259, 284097, 33657023, 29331822, 29331824, 60432289, 29331826, 29331827, 29331828, 27486282, 284508, 264509, 264905, 284907, 18108370, 66712502, 60431528, 284628, 284909, 18108372, 18108374, 56182435, 18108376, 55810764, 55811576, 35696423, 35695855, 265006, 265007, 264512, 265008, 265009, 264634, 264635, 60431850, 264636, 264555, 264592, 60431735, 284638, 33657402, 56182323, 60433356, 60433438, 264595, 55812038, 264596, 264758, 83373044, 52846317, 18108385, 33657084, 18108387, 55811386, 65274727, 56526486, 87168518, 60432113, 265017, 22279000, 265018, 265019, 284584, 18108351, 284448, 264566, 264288, 264488, 264587, 264788, 264510, 264512, 284630, 264591, 284592, 264259, 264594, 264595, 264603, 264605, 18108351, 264585, 264369, 18108354, 55812038, 265017, 284689, 35695917, 35695763, 60431528, 60432113, 22279002
2527	88094580 (5053, 5054)	Novel Protein sim. GBank gi 2258437 (AF008197) - syncollin [Rattus norvegicus]		UNCLASSIFIED		
2528	88078380 (5055, 5056)	Novel Protein sim. GBank gi 2085786 (AC002086) - similar to zinc finger 5 protein from Gallus gallus, U51840 (PID:g1399185) [Homo sapiens]	Contains protein domain (PF00551) - BTB/POZ domain	dna_ma_bind		
2529	86670926 (5057, 5058)	Novel Protein sim. GBank gi 3786433 (AF098505) - similar to Arabidopsis thaliana male sterility protein 2 (SW:Q08891) [Caenorhabditis elegans]		synthase		
2530	80259978 (5059, 5060)					284369, 284556
2531	87768931 (5061, 5062)			UNCLASSIFIED		29331822, 29331824, 60432289, 284508, 284509, 264908, 265011, 264769, 21908768, 33857023, 87168518, 22279000, 284593
2532	87419778 (5063, 5064)	Novel Protein sim. GBank gi 2864625 (emb CAA16972) - (AL021811) putative protein [Arabidopsis thaliana]				
2533	87000255 (5065, 5066)	Novel Protein sim. GBank gi 437181 (U02289) - GTPase-activating protein [Caenorhabditis elegans]		UNCLASSIFIED		284555

2534	8732322 (5087, 5088)	Novel Protein sim. GBank gij3452473 (AF084205) - serine/threonine protein kinase TAO1 [Rattus norvegicus]		UNCLASSIFIED	264259, 35696052, 264905, 265017, 21908769, 265020, 265022, 33657109, 22279000
2535	91225056 (5069, 5070)	Novel Protein sim. GBank gij4468311[emb]CAB37892] - (AL031432) dJ465N24.2.1 (PUTATIVE novel protein) (isoform 1) [Homo sapiens]			65274572, 35696286, 60432289, 29331828, 66712502, 265006, 60432229, 265017, 265018, 265019, 264288, 264369, 264689, 21908768, 265020, 265021, 264636, 60170394, 22279002
2538	94218540 (5071, 5072)	Novel Protein sim. GBank gij728836[sp]P39193[ALU6_HUMAN - III] ALU SUBFAMILY SP WARNING ENTRY IIII		kinase	18108398, 56182575, 35696286, 22278997, 22278999, 60432049, 264259, 29331824, 29331826, 29331827, 29331828, 264905, 264511, 265009, 264910, 264596, 52646317, 18108351, 264681, 264683, 18108354, 264288, 264687, 264768, 264689, 21908765, 21908766, 21908787, 265021, 52645129, 33657109, 18108374, 18108380, 56182323, 18108381, 18108388, 87168518, 60432113, 22279000, 22279002, 264567, 18108391, 65274572, 35696286, 29331822, 29331825, 29331827, 29331828, 35696052, 264908, 66712502, 264909, 265008, 265011, 264760, 264288, 264685, 35695917, 60170815, 264691, 33657023, 65274620, 33657109, 18108374, 35696423, 35695855, 264636, 264558, 60170394, 56182323, 83373044
2537	95422283 (5073, 5074)	Novel Protein sim. GBank gij4557026[re]NP_003913.1pHERC - guanine nucleotide exchange factor p532		ubiquitin	22278998, 22278999, 29331822, 29331825, 29331826, 29146499, 264908, 264112, 60170831, 87168559, 264604, 265019, 264685, 264768, 87168518, 22279000, 264565, 264586
2538	36853454 (5075, 5076)			UNCLASSIFIED	22278997, 29331828, 265008, 265009, 264758, 265010, 18108351, 264683, 264288, 21908765, 35695917, 265020, 18108374, 264587
2539	94144916 (5077, 5078)			UNCLASSIFIED	264587
2540	94218545 (5079, 5080)	Novel Protein sim. GBank gij1362647[pr]S53876 - sex-regulated protein Janus A - fruit fly (Drosophila pseudoobscura)		UNCLASSIFIED	22278997, 29331828, 265008, 265009, 264758, 265010, 18108351, 264683, 264288, 21908765, 35695917, 265020, 18108374, 264587
2541	95308238 (5081, 5082)	Novel Protein sim. GBank gij1711658[sp]P54797[T10_MOUSE - SER/THR-RICH PROTEIN T10 IN DGCR REGION		UNCLASSIFIED	264488, 264768, 264689, 264511, 20281171, 264634, 264635, 264691, 264639, 29331824, 264603, 264604, 264905, 264907, 264908, 264768

2542	95298182 (5083, 5084)	Novel Protein sim. GBank gi 5225320 gb AAD40850.1 AF08310 - (AF083107) sirinulin type 2 [Homo sapiens]	Contains protein domain (PF00220) - Neurohypophyseal hormones, N- terminal Domain	UNCLASSIFIED	264488, 18108394, 52646385, 52846842, 65274572, 22278994, 35698286, 22278996, 264259, 52645080, 29331822, 29331824, 29331827, 35698052, 33658970, 264907, 264909, 52644045, 264510, 265008, 264512, 265007, 265008, 265009, 264910, 60431735, 52646317, 52644288, 265010, 265011, 265018, 265019, 18108351, 264883, 264288, 264685, 264687, 52644229, 264789, 21908768, 21908767, 21908789, 52844150, 33657023, 33657109, 52645129, 33657182, 27486281, 27486284, 33657349, 35695763, 18108374, 35696423, 35695855, 264631, 264634, 264635, 264558, 83373044, 18108385, 18108387, 87166518, 264563, 264564
2543	94139088 (5085, 5086)	Novel Protein sim. GBank gi 5419857 emb CAB46374.1 - (AL096723) hypothetical protein [Homo sapiens]	Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)		65274572, 56182575, 22278999, 264259, 29331826, 264907, 264510, 264511, 264592, 264595, 264784, 264369, 264286, 264684, 264766, 264689, 21908765, 21908767, 21908789, 60170815, 264692, 264693, 55811578, 65274791, 264636, 264558, 18108381, 60170394, 264639, 18108385, 60432113, 22279000
2544	94218549 (5087, 5088)	Novel Protein sim. GBank gi 2498110 sp O83181 AEGP_RAT - APICAL ENDOSOMAL GLYCOPROTEIN PRECURSOR	Contains protein domain (PF00829) - MAM domain.	glycoprotein	18108397, 52646365, 22278997, 264259, 60432049, 29331822, 29331825, 29331826, 29331827, 29331828, 264905, 264908, 265006, 265007, 265008, 87166559, 265017, 265018, 265019, 18108351, 264448, 264686, 264687, 264689, 21908765, 265020, 265021, 18108370, 18108374, 18108376, 18108381, 18108385, 18108387, 56528486, 22279000, 264482, 264563, 264567
2545	87742845 (5089, 5090)	Novel Protein sim. GBank gi 3327046 dbj BAA31591 - (AB014516) KIAA0816 protein [Homo sapiens]			29331825, 264908, 265009, 60170831, 265017, 264369, 21908767, 60170815, 264692, 33657109
2546	88093861 (5091, 5092)	Novel Protein sim. GBank gi 2996032 (AF054586) - brain finger protein [Rattus norvegicus]	Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger)	UNCLASSIFIED	29331824, 265007, 22279002

2547	94143868 (5093, 5094)	Novel Protein sim. GBank gi4928607 gb AAD34064.1 AF15182 - (AF151827) CGI-69 protein [Homo sapiens]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	transport	264488, 18108394, 52646842, 18108397, 58182575, 22278995, 58894075, 22278998, 22278997, 22278999, 264259, 29331822, 29331824, 29331828, 60432289, 29331827, 35896032, 29331828, 264104, 264508, 264805, 264908, 264908, 68712502, 264909, 58182435, 285008, 285007, 264512, 285008, 265009, 60170831, 60432229, 60431735, 264594, 60433438, 21906754, 52846317, 265010, 265011, 264600, 264601, 265018, 265019, 264780, 18108351, 264682, 264448, 264288, 264369, 264684, 264688, 264687, 58181562, 284888, 264689, 21908785, 21908788, 21908767, 21908768, 29148627, 21908769, 55811957, 265020, 265021, 265022, 264690, 264691, 18108382, 264692, 264693, 27488281, 18108370, 18108374, 55810784, 55811576, 35896423, 35895855, 264635, 264638, 264555, 264637, 263981, 264557, 18108380, 264638, 58182323, 264558, 264559, 83373044, 18108385, 87188518, 22279002, 264564, 264568, 264486
2548	88179079 (5095, 5096)			UNCLASSIFIED	284488, 18108394, 52646365, 22278984, 35898286, 58994075, 22278997, 22278999, 264259, 29331822, 29331824, 29331825, 29331828, 60432289, 29331827, 29331828, 58182435, 264511, 265007, 264512, 60433356, 87188558, 264684, 264369, 52844229, 265021, 33657023, 264692, 18108374, 52844332, 264557, 18108380, 18108381, 18108382, 18108384, 18108385, 60432113, 22279000, 22279002, 264583, 264587
2549	94196883 (5097, 5098)	Novel Protein sim. GBank gi1728837 sp P39194 ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII	Contains protein domain (PF00412) - LIM domain containing proteins	struct	58182575, 22278998, 22278997, 22278998, 22278999, 264259, 264508, 264908, 29331830, 285009, 285010, 285018, 264688, 21908764, 21908765, 21908768, 21908787, 21908789, 285020, 265021, 52844150, 264691, 18108368, 60431602, 18108376, 35896423, 58182323, 18108387, 264587, 58182575, 29331822, 264105, 264512, 18108351, 35895917, 264637, 264638
2550	8778584 (5099, 5100)	Novel Protein sim. GBank gi12143888 pir J52523 - nucleoporin p82 homolog - rat (fragment)		UNCLASSIFIED	

2551	95308400 (5101, 5102)	Novel Protein sim. GBank gll4337103 gb AAD18078 - (AF129756) NG26 [Homo sapiens]	Contains protein domain (PF00561) - alpha/beta hydrolase fold	UNCLASSIFIED	18108398, 65274572, 22278995, 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 264905, 56182435, 265007, 60433438, 55812038, 21908754, 65274444, 265017, 265018, 264605, 265019, 264288, 21908768, 21908769, 21908769, 265020, 60170615, 264693, 33657109, 35696423, 264638, 56182323, 83373044, 22279000
2552	9532620 (5103, 5104)			UNCLASSIFIED	56182375, 35696286, 29331824, 29331826, 35696052, 29331828, 264508, 264907, 56182435, 265008, 264591, 33109954, 264760, 55811957, 35695917, 33657023, 33657109, 18108374, 55811576, 35696423, 35695855, 56182323, 264558
2553	95308243 (5105, 5106)	Novel Protein sim. GBank gll1711658 sp P39182 ALU5_MOUSE - SER/THR-RICH PROTEIN T10 IN DGR REGION		UNCLASSIFIED	264686, 264488, 263976, 264768, 26331826, 35698052, 35698423, 264601, 264511, 264602, 264910, 264634, 264760, 264555, 264762, 264908, 264592, 264691, 264566, 264908, 264684, 264567, 264909, 264768
2554	87761520 (5107, 5108)	Novel Protein sim. GBank gll728835 sp P39182 ALU5_HUMAN - IIII ALU SUBFAMILY SC WARNING ENTRY IIII		cadherin	22278997, 29331822, 264508, 21908769, 33657023, 33657109, 56182323
2555	87627551 (5109, 5110)	Novel Protein sim. GBank gll4884319 emb CAB3260.1 - (AL050084) hypothetical protein [Homo sapiens]		nuclease	28331824, 263972
2556	87645533 (5111, 5112)	Novel Protein sim. GBank gll4106984 (AC003038) - R30923_1 [Homo sapiens]	Contains protein domain (PF00514) - Armadillo/beta-catenin-like repeats	UNCLASSIFIED	22278998, 264508, 33657402, 264693, 264684, 264766, 264689, 33657023, 33657109, 35695855, 264558, 264567
2557	78437803 (5113, 5114)				264595
2558	87617591 (5115, 5116)	Novel Protein sim. GBank gll119110 sp P0321 EBN1_EBV - EBNA-1 NUCLEAR PROTEIN		UNCLASSIFIED	22278997, 29331824, 66714117, 29331825, 264906, 264511, 265018, 264448
2559	88096382 (5117, 5118)	Novel Protein sim. GBank gll4538998 emb CAB39619.1 - (AL049481) A/G1-like protein [Arabidopsis thaliana]		UNCLASSIFIED	22278997, 29331822, 29331828, 60433356, 265011, 264288, 264765, 264766, 264789, 21908765, 21908766, 60432113, 284482
2560	87994530 (5119, 5120)	Novel Protein sim. GBank gll5051399 emb CAB44995.1 - (AL078630) 573K1.3 (mm17M1-4 (novel 7 transmembrane receptor (rhodopsin family) (olfactory receptor LIKE protein)) [Mus musculus]	Contains protein domain (PF00001) - 7 transmembrane receptor (rhodopsin family)		
2561	88176575 (5121, 5122)	Novel Protein sim. GBank gll5328625 gb AAD2056.1 AF04495 - (AF044953) NADH:ubiquinone oxidoreductase PGIV subunit [Homo sapiens]		UNCLASSIFIED	22278995, 35696286, 22278996, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 29331826, 29331827, 29331828, 265007, 60432229, 87166559, 265017, 265018, 265019, 264689, 21908768, 21908769, 35695917, 265020, 33657023, 33657109, 18108374, 264634, 264559, 18108385, 87168518, 22279002

2562	87645539 (5123, 5124)	Novel Protein sim. GBank gij4108984 (AC003038) - R30923_1 [Homo sapiens]		UNCLASSIFIED	5894075, 22278996, 22278997, 22278999, 284259, 29331822, 60432288, 33857402, 60433356, 21908765, 55811957, 60170615, 33857023, 284893, 35695855, 87168518, 264488, 35696286, 22278999, 264259, 29331822, 29331824, 35696052, 284508, 284907, 284908, 284909, 52844045, 284510, 284511, 265009, 284910, 284591, 284593, 33837402, 265017, 265018, 265019, 18108351, 264686, 21908767, 21906768, 55811957, 35695917, 265020, 284691, 264693, 27488262, 264628, 18108374, 35696423, 35695855, 284632, 284634, 284635, 284639, 284558, 18108384, 87168518, 22279000, 22279002, 264482, 284583, 264565, 264566, 284488
2563	88095497 (5125, 5126)	Novel Protein sim. GBank gij4886447[emb]CAB43371.1] - (AL050270) hypothetical protein [Homo sapiens]		UNCLASSIFIED	29331822, 265007, 265010, 265019, 264769, 55811576, 56182323
2564	80502783 (5127, 5128)	Novel Protein sim. GBank gij1352944[sp]P47179[YJP_YEAST - HYPOTHETICAL 118.4 KD PROTEIN IN BAT2-DALS INTERGENIC REGION PRECURSOR.		sulfoltransferase	
2565	85530908 (5129, 5130)			UNCLASSIFIED	68714117, 284909, 283978, 264632
2566	80224956 (5131, 5132)	Novel Protein sim. GBank gij628012[pi]A53933 - myosin I myr 4 - rat	Contains protein domain (PF00063) - Myosin head (motor domain)	- struct	18108370, 35695855, 284558, 284558, 18108383
2567	88143580 (5133, 5134)				265020, 60170615
2568	91233099 (5135, 5136)	Novel Protein sim. GBank gij468009[sp]P34548[YNJ4_CAEEL - HYPOTHETICAL 33.8 KD PROTEIN R10E11.4 IN CHROMOSOME III		synthase	60424178, 18108394, 56181688, 56984075, 22278999, 284490, 284259, 29331822, 56182181, 29331824, 60424269, 29331825, 29331826, 29331828, 264509, 29331830, 265007, 265008, 265009, 33857402, 265010, 265011, 265019, 264448, 284683, 264288, 18108354, 264769, 21908766, 21908767, 35695917, 265021, 33657023, 18108382, 33657109, 33657182, 35695763, 60431528, 55810784, 18108379, 83373044, 18108385, 60432113, 284482

2569	95313764 (5137, 5138)	Novel Protein sim. GBank gi 2599560 gb AAB84166.1 - (AF029674) basic leucine zipper protein LZIP [Homo sapiens]	Contains protein domain (PF00170) - bZIP transcription factor	transcription factor	18108394, 56182575, 56181686, 22278995, 22278996, 56994075, 35696286, 22278997, 22278998, 22278999, 264259, 60432049, 29331822, 56182181, 29331824, 29331825, 60432289, 29331826, 29331827, 29331828, 35696052, 264905, 264906, 56182435, 265008, 265009, 264910, 60432229, 264592, 60433356, 60433436, 21906754, 87168559, 265017, 265018, 265019, 264682, 264448, 264288, 21906765, 21906766, 21906767, 21906768, 29148827, 21906769, 35695917, 265021, 265022, 52644150, 264690, 264691, 264692, 264693, 65274620, 263957, 35695763, 20281069, 263974, 18108374, 55810764, 35696423, 35695855, 264558, 18108381, 56182323, 18108382, 83373044, 18108385, 56526488, 87168518, 22279000, 22279002, 264586
2570	94138754 (5138, 5140)	Novel Protein sim. GBank gi 4758954 ref NP_004567.1 pPP2 - protein phosphatase 2 (formerly 2A), regulatory subunit B (PR 52), beta isoform	Contains protein domain (PF01240) - Protein phosphatase 2A regulatory subunit PR55	phosphatase	22278996, 29331822, 29331824, 66714117, 29331825, 60432289, 29331827, 35696052, 264907, 264510, 265007, 265009, 264758, 33109954, 265019, 264686, 264689, 265020, 265021, 264691, 264692, 264693, 18108374, 35695855, 264634, 56182323, 264639, 60170394, 83373044, 22278002, 264482, 264508, 264905, 264907, 264628, 18108351, 264555, 264556, 264557, 264558, 264559
2571	87733750 (5141, 5142)	Novel Protein sim. GBank gi 732218 sp P34609 YO60_CAEEL - HYPOTHETICAL 128.6 KD PROTEIN ZK1098.10 IN CHROMOSOME III		struct	
2572	87627560 (5143, 5144)	Novel Protein sim. GBank gi 4884319 emb CAB43260.1 - (AL050084) hypothetical protein [Homo sapiens]		UNCLASSIFIED	35696286, 29331827, 35696052, 264100, 264104, 264110, 264592, 21906754, 29148827, 29148629, 263972, 263974, 18108374, 263976, 35695855, 60170394, 264559, 18108385

2573	85313928 (5145, 5146)	Novel Protein sim. GBank gl 389138 sp P02745 C10A_HUMAN - COMPLEMENT C1Q SUBCOMPONENT, A CHAIN PRECURSOR	Contains protein domain (PF00386) - C1q domain	- complement	264488, 60424179, 65274572, 56182575, 58181868, 22278985, 58994075, 22278987, 60432049, 264259, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 29331828, 264104, 264107, 264508, 264908, 29331830, 264909, 264510, 265008, 264512, 265008, 265009, 264910, 264591, 264592, 60432228, 264593, 60433356, 264594, 60433438, 264595, 55812038, 264759, 21908754, 87188474, 265010, 265011, 87188559, 265017, 265018, 265019, 264761, 264762, 264763, 264764, 264369, 264288, 264885, 264786, 264688, 264687, 264688, 264769, 58181562, 264689, 21908765, 21908768, 21908767, 29148627, 21908768, 21908769, 265020, 265021, 265022, 60170615, 264690, 52844150, 264691, 264692, 33657023, 65274620, 18108365, 18108368, 27488265, 60431602, 284628, 60431528, 263976, 65274791, 35895855, 20281071, 60431850, 264637, 284638, 284558, 264639, 58182323, 60170394, 83373044, 18108384, 87188518, 60432113, 284482, 264564, 264585, 264566, 264567, 22278985, 284258, 60432289, 29331827, 29331828, 33656970, 264908, 265008, 264910, 264591, 33657402, 265016, 265018, 264448, 264764, 264369, 264288, 18108357, 21908765, 21908766, 21908768, 55811857, 60170815, 264691, 33657023, 284693, 33657109, 33657182, 27488261, 27486264, 33657349, 264638, 264555, 83373044, 18108385, 264482
2574	94746814 (5147, 5148)	Novel Protein sim. GBank gl 3334982 (AC005306) - R27216_1 [Homo sapiens]	Contains protein domain (PF00651) - BTB/POZ domain	UNCLASSIFIED	
2575	87754408 (5149, 5150)	Novel Protein sim. GBank gl 4928729 gb AAD34125.1 AF15188 - (AF151888) CGI-130 protein [Homo sapiens]		UNCLASSIFIED	
2576	95357881 (5151, 5152)	Novel Protein sim. GBank gl 4680681 gb AAD27730.1 AF13295 - (AF132955) CGI-21 protein [Homo sapiens]	Contains protein domain (PF00442) - Ubiquitin carboxyl-terminal hydrolases family 2	ubiquitin	264259, 35698052, 264906, 60433438, 264681, 18108351, 264288, 52644150, 264628, 35698423
2577	86996621 (5153, 5154)	Novel Protein sim. GBank gl 4337103 gb AAD18078 - (AF129756) NG26 [Homo sapiens]		UNCLASSIFIED	29331825, 265018, 265019, 264685
2578	87786841 (5155, 5156)			UNCLASSIFIED	264488, 264906, 264908, 264910, 264596, 264603, 264604, 264605, 264768, 21908769, 264628, 264630, 264634, 264639, 264583
2579	87292679 (5157, 5158)			UNCLASSIFIED	29331822, 29331824, 264767

2580	88168788 (5159, 5160)	Novel Protein sim. GBank glj2588628 (AC003080) - Similar to KIAA0299; 80% similarity to AB002297 (PID:g2224539) [Homo sapiens]			265007, 265018, 264762
2581	87899048 (5161, 5182)	Novel Protein sim. GBank glj406642 [gb]AAD20049] - (AF131809) Unknown [Homo sapiens]	Contains protein domain (PF00595) - PDZ domain (Also known as DHR or GLGF).	collagen	56994075, 29331824, 29331826, 29331828, 264905, 60433356, 60433438, 264758, 87168559, 21908769, 265022, 35695855, 263981
2582	87766789 (5163, 5164)	Novel Protein sim. GBank glj2739367 (AC002505) - putative phosphatidylinositol-4-phosphate 5-kinase [Arabidopsis thaliana]		eph	264488, 264907, 264908, 264810, 264764, 264684, 264768, 264636, 264555, 264585
2583	91220850 (5165, 5186)	Novel Protein sim. GBank glj4378112 [emb]CAA18521.1] - (AL021578) dJ453C12.2 (similar to transcription factor RBP-L) [Homo sapiens]	Contains protein domain (PF00047) - Immunoglobulin domain	transcript factor	56181886, 264259, 264510, 264512, 264591, 264592, 264593, 264594, 264595, 264598, 264603, 264628, 55810784, 264630, 264637, 264565
2584	80430841 (5167, 5168)			UNCLASSIFIED	264908, 264910, 264768, 264693, 18108374, 55811576, 58182323
2585	80438126 (5169, 5170)	Novel Protein sim. GBank glj2736151 (AF021835) - myotonic dystrophy kinase-related Cdc42-binding kinase [Rattus norvegicus]		kinase	264768
2588	91226138 (5171, 5172)				22278998, 264259, 29331822, 29331824, 29331827, 29331828, 264908, 265007, 265009, 264591, 60433358, 33637402, 265018, 264782, 264288, 21908768, 21908767, 21908769, 265022, 264691, 83373044, 58526488, 22278002
2587	80430943 (5173, 5174)				264908, 265019, 264768, 264693, 55811576, 58182323
2588	80074385 (5175, 5176)			UNCLASSIFIED	264584
2589	85515807 (5177, 5178)	Novel Protein sim. GBank glj3021598 [emb]CAA71415] - (V10389) nuclear protein [Xenopus laevis]		UNCLASSIFIED	35696052, 264905, 264906, 264907, 264908, 264909, 265009, 265018, 264769, 35698423, 264636
2590	87054526 (5179, 5180)	Novel Protein sim. GBank glj2104689 (U82793) - alpha glucosidase II, alpha subunit [Mus musculus]	Contains protein domain (PF01055) - Glycosyl hydrolases family 31	glucoamylase	22278995, 29331830, 265008, 265010, 265017, 264839
2591	94192167 (5181, 5182)	Novel Protein sim. GBank glj5702202 [gb]AAD47199.1] AF12916 - (AF129166) long-chain acyl-CoA synthetase 5 [Homo sapiens]		eph	264259, 29331822, 264106, 264906, 56182435, 265007, 265008, 33109954, 264448, 55811957, 265020, 18108370, 55811576, 22279002

2592	95332648 (5183, 5184)	Novel Protein sim. GBank gi 3024998 sp Q60936 YAB1_MOUSE - HYPOTHETICAL HEART PROTEIN		transport	18108397, 56182575, 35696286, 56994075, 264259, 29331822, 29331824, 29331826, 60432289, 29331827, 29331828, 264906, 264909, 265007, 265008, 264910, 60432229, 264594, 60433356, 60433436, 55812038, 18108348, 21906754, 265011, 87168559, 265017, 265019, 264764, 264369, 264288, 264766, 265021, 60170615, 33857023, 33657109, 264629, 35696423, 35695855, 264557, 264638, 60170394, 56182323, 83373044, 56526486, 87168518, 264583, 264482, 264585
2593	87754416 (5185, 5186)	Novel Protein sim. GBank gi 4929729 gb AAD34125.1 AF15188 - (AF15188) CGI-130 protein [Homo sapiens]		Im7	22278999, 29331825, 264758, 21906754, 52646317, 265010, 18108351, 264288, 264389, 21908768, 264693, 18108370, 264637, 264638, 264482
2594	95305758 (5187, 5188)	Novel Protein sim. GBank gi 4929587 gb AAD34054.1 AF15181 - (AF15181) CGI-59 protein [Homo sapiens]		UNCLASSIFIED	264488, 18108398, 56182575, 35696286, 22278997, 264093, 264259, 29331822, 29331825, 66714117, 29331826, 264905, 264909, 52644045, 56182435, 264510, 264512, 265007, 264757, 21908754, 87168474, 265017, 264760, 264448, 264764, 264288, 264766, 264689, 21908768, 33657109, 263975, 263977, 264634, 264556, 60170394, 56182323, 56526486, 264482, 264583, 264584, 264566, 264587
2595	79581678 (5189, 5190)			UNCLASSIFIED	264892
2596	87538637 (5191, 5192)	Novel Protein sim. GBank gi 4309681 gb AAD15478 - (AC008930) R33423_1 [Homo sapiens]		UNCLASSIFIED	22278999, 264259, 265018, 264448, 265021, 60431528
2597	94784089 (5193, 5194)			UNCLASSIFIED	264905, 264509, 264908, 264762, 264766, 35695917, 35695855, 264635, 264638, 83373044, 264488
2598	98094948 (5195, 5196)	Novel Protein sim. GBank gi 1001351 dbj BAA10838 - (D84008) hypothetical protein [Synechocystis sp.]		UNCLASSIFIED	22278998, 264259, 29331824, 87168474, 264683, 21906766, 35695917, 264691, 33657023, 33857109, 18108370, 18108374, 264564, 264565
2599	87842888 (5197, 5198)	Novel Protein sim. GBank gi 3941737 (AF109719) - BAT2 [Mus musculus]		MHC	264768, 264769, 21906768, 33657182, 35695763, 18108370, 18108374, 264635, 264636, 56528486, 22279000, 264568
2600	87787846 (5199, 5200)	Novel Protein sim. GBank gi 4263521 gb AAD153471 - (AC004044) putative WD-repeat protein [Arabidopsis thaliana]	Contains protein domain (PF00400) - WD domain, G-beta repeat	kinasereceptor	35696286, 264093, 264288, 21906769, 35696423, 35695855

2601	91243070 (5201, 5202)	Novel Protein sim. GBank gij728837 sp P39194 ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII		kinase	56182575, 22278999, 264259, 29331822, 29331824, 60432289, 29331826, 29331827, 264908, 265007, 265008, 264591, 60433356, 33657402, 60433438, 21906754, 265011, 265018, 265019, 18108351, 264448, 264369, 21906769, 265020, 60170615, 264693, 33657109, 18108370, 18108376, 58182323, 18108381, 18108385, 22278902, 264563 60433438, 21906754, 87168559, 264601, 284369, 264288, 21908787
2602	88180022 (5203, 5204)	Novel Protein sim. GBank gij4406632 gb AAD20047 - (AF131801) Unknown [Homo sapiens]			
2603	94325821 (5205, 5206)	Novel Protein sim. GBank gij3122387 sp Q81211 LIGA_MOUSE - LIGATIN		UNCLASSIFIED	264488, 65274572, 22278995, 22278998, 58994075, 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 60432289, 29331826, 29331827, 29331828, 35696052, 58182435, 264113, 265008, 265009, 60433356, 264757, 60433438, 284759, 33657084, 87168474, 265010, 285011, 87168559, 265017, 265018, 265019, 264448, 264883, 18108354, 264288, 264787, 264689, 21906765, 21906768, 21906767, 21906768, 21906769, 55811957, 265020, 265021, 265022, 60170615, 264691, 33657023, 264893, 33657109, 27486282, 18108374, 35698423, 85274791, 35698585, 264555, 264636, 264637, 58182323, 83373044, 58526486, 87168518, 60432113, 22279000
2604	94876601 (5207, 5208)	Novel Protein sim. GBank gij5454030 ref NP_006488.1 pRRP2 - RAS-related on chromosome 22		oncogene	264259, 35696032, 264508, 264906, 264907, 264908, 264909, 264510, 264512, 265008, 264910, 33657402, 264604, 264805, 264762, 264763, 264882, 264764, 264683, 264768, 264769, 264689, 33657023, 264893, 18108365, 264828, 35698423, 264831, 264632, 264634, 264635, 264637, 18108381, 264639, 83373044, 264565
2605	94316756 (5209, 5210)	Novel Protein sim. GBank gij3628745 dbj BAA33366 - (AB013721) mltaugmin 23 [Oryctolagus cuniculus]		UNCLASSIFIED	22278998, 264480, 60432049, 264259, 60432289, 264909, 265008, 60433358, 60433438, 264758, 21906754, 265010, 265011, 265018, 264681, 18108351, 264288, 264766, 264685, 21906765, 21906766, 21906768, 21906769, 264691, 264692, 264693, 85274791, 264634, 264555, 264636, 22278998, 264510, 264512, 265009, 264768, 22278902, 264566
2608	87746408 (5211, 5212)				

2607	87627742 (5213, 5214)	Novel Protein sim. GBank g 4826626 gb AAD30202.1 - (AF135022) mediator [Homo sapiens]			29331822, 29331825, 29331826, 29331827, 29331828, 264906, 264907, 264908, 66712502, 264828, 56182435, 55812038, 265010, 265017, 265018, 265019, 264768, 264688, 21908765, 55811957, 265020, 265022, 264682, 33657023, 264693, 33657109, 18108370, 264639, 56182323, 264509, 264907, 264908, 264592, 264758, 264631
2608	81734786 (5215, 5216)	Novel Protein sim. GBank g 2226005 (U9873) - ORF2: function unknown [Homo sapiens]			
2609	94843781 (5217, 5218)	Novel Protein sim. GBank g 3024889 sp P56524 Y288_HUMAN - HYPOTHETICAL PROTEIN KIAA0288 (HA6118)	Contains protein domain (PF00850) - Histone deacetylase family	histone	264488, 65274572, 35696288, 22278997, 22278999, 60432049, 264259, 56182181, 29331824, 29331825, 29331826, 60432289, 29331827, 29331828, 264905, 264907, 60433358, 60433438, 55812038, 265011, 87168559, 265017, 265018, 264448, 264765, 264288, 264766, 264689, 21908765, 21908767, 21908768, 265020, 265021, 264691, 264692, 33657109, 27488261, 18108370, 65274791, 264638, 264556, 56182323, 18108385, 56328486
2610	88177654 (5219, 5220)	Novel Protein sim. GBank g 4336855 gb AAD17989 - (AF106473) leucine-rich-domain Inter-acting protein 1; LeR inter-acting protein 1; LEAP1 [Mus musculus]		transcription factor	18108394, 22278994, 56994075, 60432049, 264259, 29331822, 29331825, 60432289, 29331827, 264107, 264109, 264905, 56182435, 264112, 265006, 265007, 265008, 265009, 60433358, 60433438, 265011, 87168559, 265017, 264448, 264682, 264764, 264288, 265021, 33657023, 263967, 33657182, 27486261, 18108374, 263976, 55811576, 264638, 87168518, 60432113
2611	87426890 (5221, 5222)	Novel Protein sim. GBank g 387676 jemb CAA92894 - (Z68760) predicted using GeneFinder; Similarity to Mouse FK508-binding protein (SW.FKB3_MOUSE) [Caenorhabditis elegans]	Contains protein domain (PF00254) - FKBP-type peptidyl-prolyl cis-trans isomerases	isomerase	22278999, 265017, 264684, 21908768, 22279000
2612	87771198 (5223, 5224)	Novel Protein sim. GBank g 5879138 gb AAD46874.1 AF16093 - (AF160934) BcDNA.LD14189 [Drosophila melanogaster]		transport	265009, 264910, 264759, 265017, 21908767, 18108385, 18108386, 60432113
2613	79481486 (5225, 5226)			UNCLASSIFIED	264685
2614	87843948 (5227, 5228)	Novel Protein sim. GBank g 5533081 gb AAD45008.1 AF16118 - (AF161181) P55T protein [Mus musculus]	Contains protein domain (PF00825) - Guanylate kinase		22278998, 22278999, 29331825, 264508, 264908, 21908754, 264602, 264766, 264769, 52844229, 21908765, 33657109, 27488264, 18108370, 263972, 264555, 60432113
2615	87381986 (5229, 5230)			UNCLASSIFIED	264768, 18108394, 264682, 264693, 264508, 264509, 264907, 264628, 264908, 264909, 18108377, 264511, 264512, 264910, 264635, 264595, 265010, 264404, 264563, 264764, 264685, 264766

2616	87428895 (5231, 5232)	Novel Protein sim. GBank glj387676 [emb CAA92994] - (Z88760) predicted using GeneFinder; Similarity to Mouse FK508-binding protein (SW:FKB3_MOUSE) [Caenorhabditis elegans]	Contains protein domain (PF00254) - FKB-type peptidyl-prolyl cis-trans isomerases	isomerase	22278995, 22278997, 22278998, 60432049, 60432289, 264828, 60433356, 264594, 60433438, 33109954, 87168474, 265011, 265017, 265019, 264288, 264768, 21908765, 21908787, 265020, 265021, 18108378, 18108377, 18108387, 87168518, 264482, 264587
2617	86976888 (5233, 5234)	Novel Protein sim. GBank glj728831 [sp P39189 ALU1_HUMAN - III ALU SUBFAMILY J WARNING ENTRY III]	kinase		265010, 265019, 264369, 264693, 55811576, 22279002
2618	91231662 (5235, 5236)	Novel Protein sim. GBank glj3319282 (AF049103) - Huntinglin Interacting protein [Homo sapiens]	Contains protein domain (PF00397) - WW domain	UNCLASSIFIED	264489, 22278996, 264490, 264259, 29331822, 264102, 264509, 264808, 264807, 68712502, 29331830, 265008, 264910, 265009, 60433356, 6043438, 264758, 21908754, 265011, 87168559, 265017, 265018, 264369, 264288, 264768, 264768, 264688, 21908765, 21908766, 21908767, 35895917, 265020, 265022, 33657023, 264692, 33657109, 264628, 18108374, 35895855, 18108381, 83373044, 18108385, 18108388, 58526488, 264563
2619	87694000 (5237, 5238)	Novel Protein sim. GBank glj2431772 (U66411) - putative type III alcohol dehydrogenase [Drosophila melanogaster]	Contains protein domain (PF00465) - Iron-containing alcohol dehydrogenases	dehydrogenase	264259, 60432288, 60433438, 21908754, 264369, 60432113, 264566
2620	95314841 (5239, 5240)	Novel Protein sim. GBank glj4322567 [gb AAD16097] - (AF090438) dachshund variant 1 [Mus musculus]	Contains protein domain (PF00628) - PHD-finger	UNCLASSIFIED	52644507, 52645156, 52646842, 65274572, 22278995, 58994075, 35896286, 22278999, 60432049, 264259, 52645080, 29331822, 29331824, 29331825, 29331826, 29331828, 35896052, 264907, 66712502, 265008, 60433356, 33657402, 52646317, 21908754, 87168474, 265010, 265017, 265018, 264448, 264369, 264288, 264687, 264768, 52644228, 264688, 264689, 21908765, 21908768, 35895917, 52644150, 264692, 33657109, 35895763, 35896423, 264556, 52644332, 18108382, 83373044, 18108385, 18108387, 65274727, 87168518, 60432113, 22279002, 264594, 264636
2621	80253485 (5241, 5242)	Novel Protein sim. GBank glj457341 [ref NP_001174.1 pATP6 - ATPase, H+ transporting, lysosomal subunit 1; vacuolar proton pump, H-ATPase subunit]			264488, 264808, 264907, 264808, 264512, 265007, 264758, 35895917, 264634, 264636, 264563, 264482
2622	81780390 (5243, 5244)	Novel Protein sim. GBank glj3880355 [emb CAB05299] - (Z82285) predicted using GeneFinder [Caenorhabditis elegans]		UNCLASSIFIED	29331824, 35896052, 265007, 265010, 264288, 29148829

2624	91639308 (5247, 5248)	Novel Protein sim. GBank gl 3860355 emb CA805299 - (Z82285) predicted using Genefinder [Caenorhabditis elegans]		UNCLASSIFIED	56181686, 22278996, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 56182181, 29331825, 60432289, 29331828, 35698052, 29146489, 66712502, 52644045, 265007, 265008, 60433356, 33109854, 21908754, 265010, 265011, 265018, 264448, 264288, 21908765, 21908766, 21908767, 29146829, 35695917, 265021, 265022, 27486285, 18108370, 60431528, 55811576, 35695855, 58182323, 18108385, 87168518, 22279002, 18108391
2625	86452068 (5249, 5250)	Novel Protein sim. GBank gl 2887429 db BAA24857 - (AB007887) KIAA0427 [Homo sapiens]		UNCLASSIFIED	284091, 284511, 263981
2626	18533787 (5251, 5252)	Novel Protein sim. GBank gl 487416 (L20302) - actin filament protein [Gallus gallus]		struct	265008
2627	87636823 (5253, 5254)	Novel Protein sim. GBank gl 89462 pir A27307 - proline-rich phosphoprotein (gene PRH1, Db allele) - human		UNCLASSIFIED	22278998, 265007, 265009, 264448, 21908767, 265021, 264558, 87168518
2628	94948254 (5255, 5256)	Novel Protein sim. GBank gl 3123552 emb CAA18609 - (AL022578) dJ393P12.2 (hypothetical Proline-rich protein KIAA0269 LIKE) [Homo sapiens]		UNCLASSIFIED	22278997, 22278999, 264259, 60432048, 29331822, 29331824, 29331825, 29331827, 35696052, 29331828, 264907, 264909, 265008, 264591, 60433356, 60433438, 265010, 265017, 265018, 264369, 264288, 18108357, 21908765, 21908768, 265022, 65274781, 264638, 18108387, 87168518, 22279002
2629	87376490 (5257, 5258)	Novel Protein sim. GBank gl 4928595 gb AAD34058.1 AF15182 - (AF151821) CGI-83 protein [Homo sapiens]		synthase	29331825, 29331826, 264102, 265006, 264768, 35695917, 264691, 33657023, 263972, 18108374, 22279000
2630	79186364 (5259, 5260)			UNCLASSIFIED	264636, 18108385
2631	84845909 (5261, 5262)	Novel Protein sim. GBank gl 321605 pir J01161 - Gag protein - Visna virus (strain EV1)	Contains protein domain (PF000098) - Zinc finger, CCHC class	dna_mn_bind	52644507, 52645156, 52646365, 52646842, 22278994, 22278995, 35698286, 59994075, 22278997, 22278999, 264259, 52645080, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 35696052, 33655970, 264905, 264509, 264907, 264908, 264511, 264512, 265007, 265008, 264910, 52646317, 33657084, 52644286, 265010, 87188559, 265017, 265018, 265019, 264760, 264762, 264448, 264288, 264369, 264768, 52644229, 21908764, 21908765, 21908766, 21908767, 21908769, 35695917, 265020, 52644150, 33657023, 52645129, 33657109, 33657182, 27486281, 27486282, 27486285, 33657348, 35695783, 35698423, 85274791, 35695855, 264634, 264637, 52644332, 59182323, 60432113, 264566, 264488
2632	36730414 (5263, 5264)				264685

2633	195011617 (5265, 5266)	Novel Protein sim. GBank gi 1139548 dbj BAA10889 - (D64009) seizure-related gene product 8 type 2 precursor (Mus musculus)	Contains protein domain (PF00084) - Sushi domain (SCR repeat)		22278995, 22278997, 22278998, 284259, 29331822, 29331824, 29331825, 29331827, 284508, 264908, 265008, 265007, 265008, 285009, 55812038, 33657084, 55811388, 265010, 265011, 87188559, 265018, 265019, 264683, 264288, 264686, 29148629, 33657023, 264693, 33657182, 35695763, 55811576, 264639, 56182323, 83373044, 18108385, 56526486, 87168518, 22279000, 22279002, 264585
2634	87330921 (5267, 5268)	Novel Protein sim. GBank gi 5441611 emb CA846854.1 - (AJ388555) hypothetical protein [Canis familiaris]	UNCLASSIFIED		29331826, 263972, 264089
2635	86623144 (5269, 5270)	Novel Protein sim. GBank gi 4680683 gb AAD27721.1 AF13294 - (AF132946) CGI-12 protein [Homo sapiens]			22278997, 264259, 29331824, 66714117, 29331827, 29331828, 284907, 33657084, 265017, 265018, 264448, 264288, 21906766, 21906767, 21906768, 29148629, 18108376, 55811576, 35695855, 87168518, 22279000
2636	87260534 (5271, 5272)	Novel Protein sim. GBank gi 3879146 emb CA807846 - (Z93386) Similarity to Yeast hypothetical 52.9 KD protein (SW:P43816); cDNA EST EMBL:M89432 comes from this gene; cDNA EST EMBL:D71008 comes from this gene; cDNA EST EMBL:D73578 comes from this gene; cDNA EST EMBL:D69025 comes ...	Contains protein domain (PF01546) - Peptidase family M20/M25/M40		264589, 29331822, 29331828, 265008, 60170831, 264681, 264765, 264685, 29148627, 21906769, 29148784, 265022, 60170615, 264635, 18108385, 56526486, 22279002, 264567
2637	95011298 (5273, 5274)	Novel Protein sim. GBank gi 4758208 ref NP_004081.1 pDUSP - dual specificity phosphatase 3 (vaccinia virus phosphatase VH11-related)	Contains protein domain (PF00782) - Dual specificity phosphatase, catalytic domain	phosphatase	284488, 264489, 52644507, 264687, 52646365, 52646842, 22278994, 22278998, 22278999, 20281171, 264259, 29331822, 52645080, 66714117, 29331825, 29331826, 29331827, 35696052, 29331828, 29146488, 284508, 264905, 284906, 284907, 284908, 264909, 264511, 264512, 264910, 264591, 264592, 60432229, 284593, 284594, 33657402, 60433356, 284757, 60433438, 284596, 264758, 52646317, 21906754, 52644296, 265010, 264600, 264602, 264603, 264605, 264781, 284782, 264681, 264448, 264784, 264765, 264288, 264786, 264686, 264788, 264687, 264789, 21906765, 21906768, 21906767, 21906768, 21906769, 35695917, 265021, 60170815, 264681, 33657023, 264692, 284693, 65274620, 27486264, 18108370, 264828, 264829, 18108374, 35698423, 35695855, 264632, 264634, 264635, 264637, 264638, 52644332, 264639, 264558, 83373044, 60432113, 264564, 264565, 264566, 264488, 264567

2636	94326733 (5275, 5276)	Novel Protein sim. GBank gi 4929689 gb AAD34105.1 AF151888 - (AF151888) CGI-110 protein [Homo sapiens]	Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	UNCLASSIFIED	60424179, 52844507, 52646842, 18108398, 58182575, 22278995, 22278996, 35698286, 22278997, 22278998, 264259, 60432049, 29331822, 60424269, 29331826, 35696052, 29146498, 264905, 52644045, 56182435, 60433356, 33657402, 55812038, 55811386, 265018, 264288, 264768, 52844228, 56181562, 28148827, 28148629, 55811957, 29148784, 35895917, 265021, 52844150, 33657023, 65274820, 33657109, 35695763, 18108374, 55810784, 35696423, 55811576, 35695855, 60431850, 56182323, 60432113, 284404
2639	95361346 (5277, 5278)	Novel Protein sim. GBank gi 2190007 dbj BAA20355 - (AB004109) phosphatidylserine synthase II [Cricetus griseus]		synthase	264488, 29331825, 35696052, 264508, 264509, 284909, 264512, 33657402, 60433438, 264758, 85658542, 264600, 285020, 265021, 33657109, 264628, 35696423, 264555, 264639, 264563, 264564, 264585, 264586, 264488
2640	87781330 (5279, 5280)	Novel Protein sim. GBank gi 3158516 (AF087617) - contains similarity to chromo (chromatin organization modifier) domains (Pfam: chromo.hmm, score: 17.76 and 27.94) and to helicases conserved C-terminal domain (Pfam: helicase_C.hmm, score: 67.00) [Caenorhabditis elegans]	Contains protein domain (PF00271) - Helicases conserved C-terminal domain	helicase	29331822, 28331826, 264908, 33109954, 265017, 265019, 21806788, 35695763, 264638, 264637, 18108387
2641	11669834 (5281, 5282)	Novel Protein sim. GBank gi 2584955 (AF030001) - unknown [Mus musculus]		UNCLASSIFIED	264828
2642	87412575 (5283, 5284)	Novel Protein sim. GBank gi 4490304 emb CAB38795.1 - (AL035878) putative protein [Arabidopsis thaliana]		UNCLASSIFIED	264259, 29331822, 33657402, 265019, 264369, 264691, 264634, 58526486, 22278002
2643	87843961 (5285, 5286)	Novel Protein sim. GBank gi 3789787 gb AAC67502.1 - (AF059568) actin binding protein MAYVEN [Homo sapiens]	Contains protein domain (PF00270) - DEAD/IDEAH box helicase	helicase	22278997, 264259, 29146499, 56182435, 264810, 265010, 18108351, 264682, 264683, 264369, 264684, 264685, 264686, 29148627, 264690, 33657109, 18108370, 263973, 18108374, 264634, 264557, 264558, 18108385, 264482
2644	88177671 (5287, 5288)	Novel Protein sim. GBank gi 3789787 gb AAC67502.1 - (AF059568) actin binding protein MAYVEN [Homo sapiens]	Contains protein domain (PF00851) - BTB/POZ domain	nuc_rept	264107, 264687
2645	17277228 (5289, 5290)	Novel Protein sim. GBank gi 11708722 sp P49749 EVX2, MOUSE - HOMEBOX EVEN SKIPPED HOMOLOG PROTEIN 2 (EVX-2)		UNCLASSIFIED	265007
2646	94148542 (5291, 5292)	Novel Protein sim. GBank gi 11708722 sp P49749 EVX2, MOUSE - HOMEBOX EVEN SKIPPED HOMOLOG PROTEIN 2 (EVX-2)		UNCLASSIFIED	264909, 264687, 264632, 83373044

2647	91212878 (5293, 5294)			UNCLASSIFIED	56182575, 22278998, 35696288, 22278998, 264259, 29331822, 56182181, 29331825, 60424269, 60432289, 35696052, 68712502, 264908, 265007, 55812038, 33109954, 21908754, 33857084, 265019, 264448, 264288, 56181582, 21908765, 21908766, 21908768, 21908769, 35695917, 265020, 285021, 52844150, 264693, 33657109, 33657349, 60431528, 18108374, 55810784, 35696423, 56182323, 60432113, 22278002, 264584
2648	87600587 (5295, 5296)				29148498, 56182435, 33109954, 265011, 264682, 55811957, 35695917, 264690, 263976, 18108377, 35696423, 60432113
2649	94128783 (5297, 5298)	Novel Protein sim. GBank gij3041852 (AC004539) - unknown function; similar to Y09105 (PID:g168617) [Homo sapiens]		UNCLASSIFIED	56182575, 35696288, 22278998, 29331824, 29331828, 60432289, 68712502, 56182435, 60170831, 60432229, 33657402, 33109954, 21908754, 265017, 264688, 264688, 21908765, 21908768, 60170815, 264693, 263987, 18108370, 263976, 60170394, 60432113, 22278002, 264563
2650	87287533 (5299, 5300)	Novel Protein sim. GBank gij5380271[dbj BAA81808.1] - (AB029335) HRPET-3 [Halocynthia roretzi]			264685
2651	88088745 (5301, 5302)	Novel Protein sim. GBank gij4240225[dbj BAA74891.1] - (AB020675) KIAA0888 protein [Homo sapiens]	Contains protein domain (PF00054) - Lamitin G domain	synthase	29331824, 29331828, 29331827, 265007, 55812038, 21908754, 18108366, 18108384, 22278002, 264587
2652	10343125 (5303, 5304)	Novel Protein sim. GBank gij4493956[emb CAB11123.2] - (Z98551) predicted using hexExon; MAL3P8.28 (PFC0845c), Hypothetical protein, len: 187 aa; Similarity to model organism hypothetical proteins (C.elegans, D.melanogaster, S.cerevisiae & S.pombe). C.elegans protein ZK287.5 (TR....		UNCLASSIFIED	264692
2653	87798735 (5305, 5306)			UNCLASSIFIED	265018, 18108370, 18108387, 264586
2654	95103240 (5307, 5308)				60424179, 65274572, 56182575, 264259, 56182181, 264908, 56182435, 55811957, 35695917, 265021, 263976, 55810784, 65274791, 56182323, 83373044, 65274727
2655	91228018 (5309, 5310)	Novel Protein sim. GBank gij3875272[emb CAB02861] - (Z87051) predicted using Genafinder; similar to Zinc finger, C3HC4 type (RING finger); cDNA EST yk443h5.3 comes from this gene; cDNA EST yk443h5.5 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger)	transcription factor	56182575, 56181888, 264092, 264259, 56182181, 60432289, 264907, 33657402, 55812038, 21908754, 87168559, 265017, 264448, 264369, 264288, 21908765, 21908766, 21908767, 21908768, 33657109, 18108370, 264628, 55811576, 264598, 264639, 83373044, 56526486, 264404, 60432113
2656	84562601 (5311, 5312)	Novel Protein sim. GBank gij3043718[dbj BAA25523] - (AB011189) KIAA0597 protein [Homo sapiens]			264693

2657	52561728 (5313, 5314)	Novel Protein sim. GBank gi15889509[dbjBAA83038.1] - (AB029009) KIAA1088 protein [Homo sapiens]			dna_rna_bind	264693
2658	80682454 (5315, 5316)	Novel Protein sim. GBank gi13688089 (AC005757) - R32611_1 [Homo sapiens]		Contains protein domain (PF00560) - Leucine Rich Repeat	nucleaseinhib	35686286, 264259, 29331822, 29331824, 29331828, 29331828, 265019, 264683, 21906788, 35695917, 264693, 35695855, 264637, 87168518, 264486, 264567
2659	87600755 (5317, 5318)	Novel Protein sim. GBank gi15420387[embjCAB46678.1] - (AJ243459) proteophosphoglycan [Leishmania major]		Contains protein domain (PF01426) - BAH domain	UNCLASSIFIED	264909, 264910, 265018, 264389, 264769, 21906789, 264693, 263972, 18108388
2660	91718472 (5319, 5320)	Novel Protein sim. GBank gi1728837[spP39194]ALU7_HUMAN - III ALU SUBFAMILY SQ WARNING ENTRY III		Contains protein domain (PF00038) - EF hand	kinase	264486, 65274572, 35686286, 22278988, 22278989, 264259, 29331822, 29331824, 60432289, 29331828, 35698052, 264908, 56182435, 265008, 265009, 60433356, 264594, 265010, 265018, 55811150, 18108351, 264682, 264684, 264389, 264288, 264687, 21906765, 29148784, 35695917, 60170615, 52644150, 33657023, 33657109, 35698423, 35695855, 264556, 60170394, 18108365, 22279000, 22279002
2661	95342817 (5321, 5322)	Novel Protein sim. GBank gi14758048[refNP_004739.1]pCPR8 - cell cycle progression 8 protein			glycoprotein	60432049, 264259, 29331824, 29331825, 29331826, 29331827, 29331828, 264906, 264909, 264593, 33109954, 265010, 265017, 265018, 265019, 264760, 264448, 264369, 264288, 21906765, 21906768, 265022, 264691, 33657023, 27488282, 60431528, 18108374, 35695855, 18108388, 264482, 264555, 264556, 264558, 264486
2662	80228739 (5323, 5324)	Novel Protein sim. GBank gi13874714[embjCAA91263] - (Z66494) similar to choline dehydrogenase: cDNA EST yk346d5.5 comes from this gene; cDNA EST yk346d5.3 comes from this gene [Caenorhabditis elegans]			dehydrogenase	264906, 264909, 264757, 264758, 264767, 264691, 33657023, 264638
2663	87780623 (5325, 5326)	Novel Protein sim. GBank gi1388670 (U58977) - Notch homolog Scalloped wings [Lucilia cuprina]		Contains protein domain (PF00008) - EGF-like domain	oncogene	35696286, 264509, 264595, 264288, 264685, 264686
2664	85518329 (5327, 5328)	Novel Protein sim. GBank gi14884408[embjCAB43311.1] - (AL050190) hypothetical protein [Homo sapiens]			UNCLASSIFIED	35696286, 22278989, 29331822, 35696052, 264906, 264907, 264909, 264510, 264511, 264512, 264593, 60433438, 265019, 264681, 21906765, 21906768, 21906767, 21906768, 265020, 265022, 35696423, 35695855, 22279002, 264482, 264486
2665	87770682 (5329, 5330)	Novel Protein sim. GBank gi15108956[gb AAD39806.1]AF113615 - (AF113615) FH1/FH2 domain-containing protein FHOS [Homo sapiens]			UNCLASSIFIED	29331825, 265007, 264910, 60432229, 265019, 264288, 21906767, 264558, 22279002
2666	87826472 (5331, 5332)	Novel Protein sim. GBank gi12500570[sp Q17533]RNP_H_CAEEL - RIBONUCLEASE PH-LIKE PROTEIN B0564.1		Contains protein domain (PF01138) - nuclease 3' exoribonuclease family	nuclease	264907, 29331830, 264681, 264683, 264288, 35695855, 264632, 264558, 264557, 264558, 264559, 264563, 264585, 264587
2667	87422720 (5333, 5334)	Novel Protein sim. GBank gi12500570[sp Q17533]RNP_H_CAEEL - RIBONUCLEASE PH-LIKE PROTEIN B0564.1				

2668	91216716 (5335, 5336)	Novel Protein sim. GBank gi 5454186 ref NP_006327.1 pZYG - ZYG homolog		UNCLASSIFIED	56181686, 35696286, 22278998, 22278999, 56182181, 29331824, 60424289, 29331825, 35696052, 29331828, 68712502, 58182435, 60433356, 284758, 21908754, 55811386, 265011, 87168559, 265017, 265019, 55811150, 284448, 284369, 284288, 21906765, 21908768, 21906768, 55811957, 35695917, 265020, 265021, 33657023, 284692, 33857109, 35695763, 80431528, 18108374, 35696423, 55811576, 284634, 80431850, 83373044, 18108385, 87188518, 22278000, 284583, 284564
2669	95415721 (5337, 5338)	Novel Protein sim. GBank gi 2147012 pir JC4899 - proline rich protein - rat			284489, 264689, 21906767, 65274572, 56102575, 21906768, 29148827, 21906769, 29148629, 35696286, 35695917, 22278998, 22278998, 265021, 265022, 60170815, 52644150, 60432048, 284259, 284691, 33657023, 284692, 29331822, 29331824, 29331825, 60432289, 33657109, 29331828, 29331827, 35696032, 29331828, 29148498, 29148489, 284905, 264908, 52644045, 284909, 56182435, 35696423, 65274791, 35695855, 265006, 284910, 284635, 60432229, 284592, 264638, 56182323, 60433356, 60170394, 60433438, 284559, 264595, 55812038, 33109954, 87168559, 60432113, 265019, 264448, 284369, 284684, 264288
2670	87613234 (5339, 5340)	Novel Protein sim. GBank gi 1723523 sp Q10382 YD8B_SCHPO - HYPOTHETICAL 94.9 KD PROTEIN C22E12.11C IN CHROMOSOME I	Contains protein domain (PF00628) - PHD-finger		18108370, 263974
2671	91214936 (5341, 5342)	Novel Protein sim. GBank gi 4768277 gb AAD2944.1 AF06425 - (AF064255) very long-chain acyl-CoA synthetase homolog 2; VLCS-H2 [Homo sapiens]	transport		52646842, 56994075, 264259, 29331822, 29331824, 29331825, 29331827, 33658970, 284509, 265006, 33109954, 21906754, 264682, 284288, 265021, 33657023, 33657109, 33657182, 27486261, 27486282, 27486285, 18108376, 18108385
2672	87399123 (5343, 5344)	Novel Protein sim. GBank gi 4966346 gb AAD34677.1 AC00634 - (AC006341) Contains two PF01344 Kelch motif domains. [Arabidopsis thaliana]	Contains protein domain (PF01344) - Kelch motif	UNCLASSIFIED	284767

2673	87430749 (5345, 5348)	Novel Protein sim. GBank gij5457337[emb CAB41505.2] - (AJ236876) poly(ADP-ribose) polymerase-2 (Homo sapiens)	Contains protein domain (PF00644) - Poly(ADP-ribose) polymerase catalytic region.	polymerase	22278995, 22278996, 22278997, 22278999, 29331822, 29331824, 29331826, 35698052, 60433436, 87168474, 87168559, 265017, 265018, 265019, 264448, 21908768, 21908769, 265020, 265021, 33657109, 27488282, 35895763, 60431850, 60170394, 87168518, 264583
2674	94847721 (5347, 5348)	Novel Protein sim. GBank gij4758824[ref NP_004280.1 pNRF3 - nuclear factor (erythroid-derived 2)-like 3	Contains protein domain (PF00170) - bZIP transcription factor	transcription factor	264488, 22278996, 35698286, 264091, 264259, 29331824, 29331826, 35698052, 264511, 55812038, 85658542, 264768, 21908765, 35895917, 264629, 35698423, 18108393, 87168518
2675	79563835 (5349, 5350)			UNCLASSIFIED	264691
2676	79628393 (5351, 5352)			UNCLASSIFIED	264908, 265008
2677	84328600 (5353, 5354)	Novel Protein sim. GBank gij1078042[pil S32154 - acetyl-CoA synthetase - fruit fly (Drosophila melanogaster)	Contains protein domain (PF00501) - AMP-binding enzyme	synthase	18108394, 65274572, 56182575, 22278994, 22278995, 56994075, 22278996, 35698286, 22278997, 22278999, 264490, 264259, 52845080, 29331824, 29331825, 60432288, 29331827, 29331828, 35698052, 29148499, 29331830, 264908, 52644045, 265006, 265007, 265008, 265009, 60432228, 60433356, 60433438, 55812038, 265010, 265011, 87168559, 265017, 265019, 18108351, 264682, 264448, 264683, 264288, 21908765, 21908766, 21908767, 21908768, 21908769, 55811957, 265020, 265021, 60170815, 52644150, 264691, 33657023, 263967, 33657109, 27488284, 27488285, 33657349, 35695763, 18108370, 18108374, 18108377, 55811576, 35698423, 35695855, 83373044, 18108387, 22279000, 22279002, 264584

2678	95001694 (5355, 5356)	Novel Protein sim. GBank gl 96760 pir A0465 - alanine transaminase (EC 2.6.1.2), cytosolic - human		UNCLASSIFIED	264488, 263994, 264489, 18108394, 52846842, 35698286, 22278999, 264259, 29331825, 35698052, 264508, 264905, 264509, 264908, 264907, 264908, 264909, 264510, 264511, 265008, 264512, 265007, 265008, 264910, 265009, 264593, 264594, 264598, 264758, 33109954, 52646317, 21906754, 265010, 265011, 87168559, 264600, 264601, 264602, 264603, 265017, 265019, 264605, 264760, 264762, 18108351, 264763, 264682, 264683, 264764, 264288, 264768, 264687, 264768, 264769, 264689, 21906765, 21908767, 35695917, 265020, 265021, 52644150, 264691, 33657023, 264693, 18108384, 18108365, 33657109, 33857349, 264628, 18108374, 35698423, 35695855, 264630, 264631, 264632, 264634, 264635, 264555, 264636, 264637, 264638, 264639, 18108385, 264563, 264564, 264565, 264566, 264488, 264567
2679	95361544 (5357, 5358)	Novel Protein sim. GBank gl 1709233 sp P07514 NC5R_BOVIN - NADH-CYTOCHROME B5 REDUCTASE	Contains protein domain (PF00970) - reductase FAD/NAD-binding Cytochrome reductase		264488, 22278996, 35696286, 264259, 29331826, 29331827, 29331828, 264909, 58182435, 264113, 264511, 265008, 60433438, 264758, 85658542, 87168474, 265011, 265017, 265019, 264288, 21908766, 21908767, 21908768, 21908769, 55811957, 265021, 264690, 33657023, 55810784, 35698423, 55811576, 264631, 18108381, 60170394, 83373044, 87168518, 264568
2680	87800356 (5359, 5360)	Novel Protein sim. GBank gl 4589604 db BAA76824.1 - (AB023197) KIAA0980 protein [Homo sapiens]	Contains protein domain (PF00036) - EF hand	struct	264259, 264102, 264905, 264908, 265007, 265008, 33109954, 265011, 18108351, 264768, 33657023, 20281149, 263972, 264630, 264635, 264638
2681	80933844 (5361, 5362)	Novel Protein sim. GBank gl 728837 sp P39194 ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII		kinase	264489, 22278995, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 265008, 264512, 264910, 264594, 60433438, 264758, 264603, 264604, 265019, 264605, 264760, 264764, 264687, 264768, 264769, 21908769, 35695917, 265020, 33657023, 264631, 264635, 264637, 264638, 264639, 264486
2682	94138934 (5363, 5364)	Novel Protein sim. GBank gl 423468 pir JQ1974 - HTF9-C protein - mouse	Contains protein domain (PF00078) - RNA recognition motif, (a.k.a. RRM, RBD, or RNP domain)	UNCLASSIFIED	35698286, 264908, 55811386, 265017, 55811150, 55811957, 35695917, 60431528, 55810764, 55811576, 35698423, 65274791, 56528488
2683	87774405 (5365, 5366)	Novel Protein sim. GBank gl 5114351 gb AAD40286.1 - (AF158271) RING finger protein terf [Homo sapiens]	Contains protein domain (PF00822) - SPRY domain	interleukinrecept	264909, 264769, 264635, 264636

2684	85787151 (5367, 5368)	Novel Protein sim. GBank gij4888469jemb CA843385.1 - (AL050284) hypothetical protein [Homo sapiens]			264593	
2685	88054289 (5369, 5370)	Novel Protein sim. GBank gij3342729 (AC005331) - R31341.2 [Homo sapiens]		UNCLASSIFIED		
2686	87628680 (5371, 5372)	Novel Protein sim. GBank gij4650844jcb BAA77027.1 - (AB026190) Kelch motif containing protein [Homo sapiens]	Contains protein domain (PF01344) - struct Kelch motif		264259, 29331822, 60432289, 29331827, 29331830, 264909, 264512, 264598, 264769, 264534, 264555, 264556, 264557, 264558, 80170394, 264559, 264488	
2687	87898183 (5373, 5374)	Novel Protein sim. GBank gij5281314jgb AAD41475.1 AF133123 - (AF133123) transcription factor IIC102 [Homo sapiens]	Contains protein domain (PF00515) - transcriptfactor TPR Domain		18108394, 18108396, 22278996, 3569286, 22278997, 29331826, 29331828, 68712502, 21908754, 265011, 264760, 264761, 264763, 264689, 21908765, 35696423, 264559, 18108385, 264563	
2688	79059584 (5375, 5376)				264908, 264760	
2689	84122440 (5377, 5378)	Novel Protein sim. GBank gij3880023jemb CAA97339 - (Z73098) Similarity to yeast hypothetical protein (Swiss Prot accession number Q09695); cDNA EST EMBL:D72982 comes from this gene; cDNA EST EMBL:D75947 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00560) - Leucine Rich Repeat		22278997, 22278998, 22278999, 29331824, 35696052, 264908, 264909, 56182435, 264512, 264910, 265009, 60433438, 21908754, 18108351, 264682, 264683, 264767, 21908765, 21906766, 21906768, 33657023, 33657182, 27486262, 27486264, 33657349, 18108370, 18108372, 18108374, 35696423, 35695855, 18108385, 22278002	
2690	88003055 (5379, 5380)	Novel Protein sim. GBank gij2477513 (AC002398) - F25965.3 [Homo sapiens]	Contains protein domain (PF00820) - struct RhoGAP domain			
2691	91219241 (5381, 5382)	Novel Protein sim. GBank gij4107276jemb CAA67130 - (X98508) acetyl-CoA synthetase [Solanum tuberosum]	synthase		65274572, 56182575, 264259, 29331822, 29331824, 29331826, 29331828, 264112, 265009, 55812038, 264596, 33109954, 265017, 264448, 264288, 264369, 264684, 21908769, 80170615, 60431528, 55810764, 264634, 264638, 264556, 264637, 22279002, 264594, 264566	
2692	84111914 (5383, 5384)	Novel Protein sim. GBank gij3513303 (AC005594) - R26984.1 [Homo sapiens]	Contains protein domain (PF00328) - peptidase Prolyl oligopeptidase family			
2693	20438807 (5385, 5386)			UNCLASSIFIED	264582	
2694	84111918 (5387, 5388)	Novel Protein sim. GBank gij3122400 sp O35682 MUG_MOUSE - MYELOID UPREGULATED PROTEIN			264559	
2695	85345513 (5389, 5390)	Novel Protein sim. GBank gij4972740jgb AAD34765.1 - (AF132177) unknown [Drosophila melanogaster]	collagen		35696286, 56994075, 22278999, 264259, 35696052, 29331830, 265011, 264288, 56181562, 264680, 264692, 33657023, 27486262, 263976, 18108376, 35696423, 35695855, 80170394, 83373044, 56526486, 22279000, 22279002, 264566	
2696	87874040 (5391, 5392)	Novel Protein sim. GBank gij728831 sp P39188 ALU1_HUMAN - IIII ALU SUBFAMILY J WARNING ENTRY IIII	synthase		264594, 21908768, 18108370, 18108372	

2697	91638472 (5393, 5394)	Novel Protein sim. GBank gij5689473dbjBAA83020.1) - (AB028991) KIAA1068 protein [Homo sapiens]		UNCLASSIFIED	35696286, 264258, 29331824, 29331825, 29331826, 29331827, 29331828, 35696052, 265006, 265007, 264512, 265009, 60170831, 60433356, 264595, 284758, 87168474, 265010, 265011, 87168559, 264601, 265017, 265018, 265019, 264761, 18108351, 264448, 264288, 264768, 264888, 264889, 21906765, 21906768, 265020, 265021, 60170815, 33657109, 18108378, 35698423, 35698585, 264555, 264558, 60170394, 264558, 18108387, 56528486, 87168518, 22278002, 264583, 264482
2698	94325891 (5395, 5396)	Novel Protein sim. GBank gij841318 (U22818) - mutant sterol regulatory element binding protein-2 [Cricetulus griseus]	Contains protein domain (PF00412) - LIM domain containing proteins	transcriptfactor	22278995, 35696288, 22278996, 22278997, 22278998, 22278999, 264258, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 33656970, 264906, 29331830, 284909, 52844045, 264910, 60433358, 33657402, 33109854, 265017, 265018, 265019, 264288, 21906765, 21906768, 21906769, 29148629, 35695917, 265021, 265022, 52644150, 33657023, 33657182, 27486281, 35698423, 65274791, 264838, 60432113, 22278000
2699	87780650 (5397, 5398)			UNCLASSIFIED	264788, 18108357, 264690, 264691
2700	94139836 (5399, 5400)	Novel Protein sim. GBank gij5174395jrefNP_006006.1pB120 - Brain protein 120		UNCLASSIFIED	60424269, 56182435, 60432229, 60433438, 55811386, 265017, 55811150, 264448, 56181582, 55811957, 264693, 33657109, 60431528, 264629, 55810784, 55811578, 65274791, 60431850, 60432113
2701	94148584 (5401, 5402)	Novel Protein sim. GBank gij1572801 (U70854) - F38A5.1 gene product [Caenorhabditis elegans]			18108394, 52845158, 35696286, 264259, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 35696052, 29331828, 29148499, 265006, 60433356, 33657402, 60433438, 264595, 33657084, 18108351, 264288, 264769, 18108359, 21906768, 35695917, 33657023, 27486261, 18108374, 18108379, 35698423, 18108382, 83373044, 18108384, 18108388, 60432113, 22278000
2702	57285368 (5403, 5404)	Novel Protein sim. GBank gij2605967 (AF030027) - 24 [Equine herpesvirus 4]		UNCLASSIFIED	29331828, 264512, 264555, 264558, 264557, 264558, 264559
2703	87849514 (5405, 5406)	Novel Protein sim. GBank gij5689399jdbjBAA82983.1) - (AB028954) KIAA1031 protein [Homo sapiens]	Contains protein domain (PF00642) - Zinc finger C-x8-C-x5-C-x3-H type (and similar).		60432289, 265007, 21906765, 21906768, 265021, 264563

2704	87649515 (5407, 5408)	Novel Protein sim. GBank gij4335894[gbjAAB63294] - (AF008554) Implantation-associated protein [Rattus norvegicus]			264488, 22278995, 22278998, 29331828, 29148499, 264905, 264908, 264907, 52844045, 264511, 33857402, 284600, 264602, 265017, 264605, 264781, 18108351, 264784, 264687, 264769, 265021, 264691, 264692, 18108382, 264693, 18108370, 18108374, 264634, 264635
2705	87771745 (5408, 5410)				264489, 264509, 264511, 264512, 264910, 264593, 87168474, 264604, 264288, 264687, 264769, 264638, 264568, 264488
2706	94326789 (5411, 5412)	Novel Protein sim. GBank gij3255952[embjCAA16821.1] - (AL021728) /prediction=(method: fmatch=(desc: [Drosophila melanogaster])	UNCLASSIFIED		264488, 52646842, 65274572, 22278994, 56994075, 22278997, 264259, 29331824, 29331825, 29331826, 29331828, 33856970, 264907, 264908, 264909, 52844045, 56182435, 265008, 265007, 60433438, 55812038, 21908754, 52644298, 265010, 264601, 265017, 265019, 264681, 264448, 264682, 264288, 264688, 264687, 264688, 21908768, 21908769, 55811937, 35695917, 265020, 265021, 60170815, 264690, 264691, 33857023, 264692, 264693, 65274620, 27488264, 263972, 18108374, 18108377, 264635, 264636, 264558, 60170394, 83373044, 65274727, 87168518, 22279000, 22278998, 22278999, 56182435, 21908754, 87168559, 265017, 264448, 52645129, 65274572, 264259, 29331822, 29331825, 60432289, 29331826, 29331827, 29331828, 264909, 264510, 265007, 264910, 60433356, 60433438, 33109954, 265010, 265011, 264369, 264288, 264765, 264693, 264565, 29331822, 18108370, 18108374, 83373044
2707	88089839 (5413, 5414)	Novel Protein sim. GBank gij3417294 (AC004381) - Unknown gene product [Homo sapiens]			29331824, 264759, 264693, 18108382, 18108388
2708	91011351 (5415, 5416)	Novel Protein sim. GBank gij545790[bbsj147178 - DARPP-32=dopamine and cAMP-regulated phosphoprotein (human, brain, Peptide, 204 aa]	UNCLASSIFIED		
2709	84853988 (5417, 5418)	Novel Protein sim. GBank gij3169705 (AC004780) - F17127.1 [Homo sapiens]	UNCLASSIFIED		
2710	87627879 (5419, 5420)	Novel Protein sim. GBank gij4468311[embjCAB37892] - (AL031432) dJ465N24.2.1 (PUTATIVE novel protein) (isoform 1) [Homo sapiens]			

2711	94111820 (5421, 5422)	Novel Protein sim. GBank g j3122400 sp O35882 MUG_MOUSE - MYELOID UPREGULATED PROTEIN		UNCLASSIFIED	284488, 264687, 52645156, 264769, 21908784, 21908765, 21908767, 21908768, 21908769, 55811957, 56994075, 22278997, 22278998, 265020, 265021, 264690, 264259, 264691, 264692, 33657023, 29331822, 29331824, 60424269, 29331826, 33657182, 29331827, 27486262, 33657349, 264508, 264905, 264907, 60431528, 264908, 264909, 55810764, 35696423, 65274791, 35695855, 265007, 264910, 60431850, 60432229, 284557, 284558, 55812038, 33108954, 18108385, 21908754, 33657084, 87168518, 87168474, 87168559, 60432113, 265017, 22279000, 265018, 265019, 22279002, 264760, 55811150, 264681, 18108351, 264565, 264764, 264566, 264288, 264768
2712	94312071 (5423, 5424)	Novel Protein sim. GBank g 5081315 gb AAD39343.1 AF07660 - (AF076607) prediabetic NOD sera-reactive autoantigen [Mus musculus]	Contains protein domain (PF00515) - TPR Domain	transferase	284488, 35696288, 22278998, 284259, 29331824, 60432289, 35696052, 264508, 284908, 66712502, 52844045, 265008, 60432229, 33657402, 60433358, 285010, 285019, 18108351, 264681, 264288, 264685, 21908765, 21908768, 21908769, 21908789, 55811957, 35695917, 265020, 265021, 60170615, 264691, 264692, 33657023, 33657108, 33657182, 33657349, 18108370, 18108374, 35696423, 35695855, 284555, 52844332, 56182323, 87168518, 60432113 66714117, 264906, 264563
2713	88003064 (5425, 5426)	Novel Protein sim. GBank g 2477513 (AC002398) - F25965_3 [Homo sapiens]		UNCLASSIFIED	264638
2714	13528218 (5427, 5428)			UNCLASSIFIED	264508, 264905, 264907, 264908, 264909, 284910, 264758, 265011, 264760, 264764, 284288, 264768, 264768, 284636
2715	94122454 (5429, 5430)	Novel Protein sim. GBank g 4321968 gb AAD15897 - (AF067430) Smarce1-related protein [Mus musculus]		UNCLASSIFIED	
2716	88003068 (5431, 5432)	Novel Protein sim. GBank g 2477513 (AC002398) - F25965_3 [Homo sapiens]		glycoprotein	284091, 264259, 29331822, 66714117, 284908, 264369, 264693, 264556, 264563 284593, 264558
2717	80077461 (5433, 5434)	Novel Protein sim. GBank g 3327046 gb BAA31591 - (ABD14516) KIAA0816 protein [Homo sapiens]			
2718	78604062 (5435, 5436)				284693
2719	68180423 (5437, 5438)	Novel Protein sim. GBank g 746485 (U23515) - weakly similar to gastrula zinc finger protein [Caenorhabditis elegans]		UNCLASSIFIED	29331822, 87168559, 265019, 265021, 52844150, 264691

2720	95086242 (5439, 5440)	Novel Protein sim. GBank gl 1335673 (U46690) - ATP-dependent RNA helicase [Mus musculus]	Contains protein domain (PF00270) - helicase DEAD/DEAH box helicase	18108374, 60424179, 264489, 56182435, 21908765, 21908768, 35698423, 22278997, 265020, 265022, 265006, 265008, 264082, 264636, 60432228, 264691, 264692, 33657023, 264693, 33657402, 83373044, 29331824, 18108368, 60424269, 29331826, 18108385, 52645129, 21908754, 35698052, 29331828, 87188474, 264100, 265010, 265011, 265019, 22279002, 264905, 264482, 264563, 264906, 18108351, 264681, 18108370, 29331830, 264908, 68712502, 52644045, 264909, 264828, 18108354, 22278995, 35698286, 264259, 29331822, 29331824, 60714117, 29331826, 264906, 60433436, 265017, 18108351, 264448, 264288, 264769, 21908768, 265021, 33657109, 263989, 80431528, 264629, 55811578, 65274791, 35695855, 264631, 264637, 60170394, 58182323, 22279000, 35698286, 22278999, 21908754, 265017, 264762, 264288, 21908765, 21908767, 21908768, 35695917, 18108362, 27486282, 35695855, 264558, 264559
2721	95345523 (5441, 5442)	Novel Protein sim. GBank gl 4828663 gb AAD34092.1 AF15185 - (AF151855) CGI-97 protein [Homo sapiens]	Contains protein domain (PF01172) - Uncharacterized protein family UPF0023	264508, 264509, 264808, 264909, 264910, 55812036, 264766, 264687, 264629, 264636, 264488
2722	91638807 (5443, 5444)	Novel Protein sim. GBank gl 3212997 gb AAC23434.1 - (AC004897) match to ESTs AA667999 (NID:g2826700), AA165465 (NID:g1741481), Z45871 (NID:g575105), and T84028 (NID:g712314); similar to various tre-like proteins including: AF040054 (PID:g2746883), D13644 (PID:g2104571), AL02114...	Contains protein domain (PF00566) - oncogene TBC domain	18108398, 22278999, 20281099, 29331824, 29331826, 60432228, 29331828, 80170831, 60432228, 60433438, 18108351, 264682, 21908766, 21908767, 21908769, 35695917, 33657023, 33657109, 18108372, 18108374, 35695855, 22279000, 22279002
2723	87387732 (5445, 5446)		UNCLASSIFIED	264488, 52644507, 264259, 29331827, 21908754, 265011, 18108351, 264448, 264288, 264685, 264689, 35695917, 265020, 33657182, 27486281, 18108370, 18108374, 35698423, 18108385, 22279000
2724	87638563 (5447, 5448)	Novel Protein sim. GBank gl 4880881 gb AAD27730.1 AF13295 - (AF132855) CGI-21 protein [Homo sapiens]	ubiquitin	264488, 264828, 264685
2725	94853991 (5449, 5450)	Novel Protein sim. GBank gl 3169705 (AC004780) - F17127_1 [Homo sapiens]	UNCLASSIFIED	
2726	86880599 (5451, 5452)	Novel Protein sim. GBank gl 3342738 (AC005328) - R28660_1, partial CDS [Homo sapiens]	MHC	

2727	81010470 (5453, 5454)	Novel Protein sim. GBank gi 7312671sp p39219 RLUA_ECOLI - RIBOSOMAL LARGE SUBUNIT PSEUDOURIDINE SYNTHASE A (PSEUDOURIDYLATE SYNTHASE) (URACIL HYDROLYASE)	Contains protein domain (PF00849) - RNA pseudouridylylate synthase	- synthase	65274572, 56182575, 22278994, 56984075, 22278996, 22278997, 22278999, 60432049, 29331822, 29331824, 66714117, 29331828, 29331827, 35696052, 29331828, 33656970, 264508, 66712502, 264910, 33657402, 60433438, 264758, 55812038, 21908754, 33657084, 55811386, 265018, 265019, 264767, 21908765, 21908767, 21908769, 55811957, 35695917, 52644150, 33657023, 33657109, 33657182, 27486281, 27486282, 27486285, 33657349, 55811576, 35696423, 35695855, 264630, 60431850, 264638, 56182323, 67168518, 60432113, 22279000, 264564, 264583
2728	94126022 (5455, 5456)	Novel Protein sim. GBank gi 3880433 emb CAA91399 - (Z66521) similar to mitochondrial RNA splicing MSR4 like protein; cDNA EST EMBL:CD9217 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	UNCLASSIFIED	18108394, 56182435, 21908767, 55811857, 35695855, 265021, 264690, 264596, 264259, 264557, 29331822, 264559, 264448, 264288
2729	94126024 (5457, 5458)	Novel Protein sim. GBank gi 2408095 emb CAB16300 - (Z99168) putative RNA splicing protein [Schizosaccharomyces pombe]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	UNCLASSIFIED	65274572, 264259, 60432289, 66712502, 56182435, 264448, 264288, 264389, 55811957, 265021, 264557, 60432113
2730	94126026 (5459, 5460)	Novel Protein sim. GBank gi 3880433 emb CAA91399 - (Z66521) similar to mitochondrial RNA splicing MSR4 like protein; cDNA EST EMBL:CD9217 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	transport	264887, 29331824, 29331826, 35696052, 264107, 56182435, 265008, 265009, 264592, 60431735, 265011, 264601, 265017, 18108351, 264288, 29148627, 55811957, 265021, 264690, 18108368, 18108374, 264557, 264558, 264559, 18108387, 56528488, 264566, 264486
2731	87723022 (5481, 5482)	Novel Protein sim. GBank gi 1723239 sp Q10166 YAUB_SCHPO - HYPOTHETICAL 35.7 KD PROTEIN C26A3.11 IN CHROMOSOME I	Contains protein domain (PF00795) - Carbon-nitrogen hydrolase		264259, 35696052, 265008, 264758, 264762, 264448, 264288, 29148627, 21908769, 87188518, 22279002
2732	94126028 (5463, 5464)	Novel Protein sim. GBank gi 3880433 emb CAA91399 - (Z66521) similar to mitochondrial RNA splicing MSR4 like protein; cDNA EST EMBL:CD9217 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	UNCLASSIFIED	264887, 264489, 18108358, 56182435, 264689, 35696423, 55811957, 265021, 265008, 265008, 264910, 265009, 264690, 264555, 264259, 264556, 264557, 264558, 264559, 18108383, 33657109, 87168518, 265010, 264601, 60432113, 265017, 264905, 264448, 263972, 264369, 264587
2733	87363060 (5465, 5466)			UNCLASSIFIED	29331825, 264509, 264909
2734	94140286 (5467, 5468)	Novel Protein sim. GBank gi 4519621 dbj BAA75670.1 - (AB017614) OASIS protein [Mus musculus]	Contains protein domain (PF00170) - bZIP transcription factor	transcription factor	60424179, 52644507, 56182575, 264259, 29331828, 264907, 264510, 264910, 60433356, 265019, 55811150, 264681, 264763, 264687, 33657182, 18108370, 60431528, 60431850, 56182323, 83373044

2735	87712338 (5469, 5470)	Novel Protein sim. GBank gij3850569 (AC005279) - ESTs gbiT21278, gbiT45403, and gbiTAA586113 come from this gene. [Arabidopsis thaliana]		glycoprotein	22278988, 60432289, 29331827, 29148498, 284108, 264909, 264112, 33657402, 87168474, 285017, 264762, 264448, 264764, 264684, 21908765, 264693, 33657109, 263976, 264636, 264638, 264557, 22278000, 22278002, 264567
2736	80247655 (5471, 5472)			UNCLASSIFIED	264905, 264628, 264628, 263978, 264632, 284584
2737	87604526 (5473, 5474)				264680
2738	85731808 (5475, 5476)	Novel Protein sim. GBank gij2558501 [dbjBAA22898] - (D63850) hepatoma-derived growth factor [Mus musculus]			264488, 265009, 264768, 264691
2739	94319834 (5477, 5478)	Novel Protein sim. GBank gij5420387 [embjCAB46679.1] - (AJ243459) proteophosphoglycan [Leishmania major]		UNCLASSIFIED	264694, 83373044, 264568
2740	94148762 (5479, 5480)	Novel Protein sim. GBank gij3417386 [embjCAA75495] - (Y15197) microtubule-associated protein, MAP-115 [Mus musculus]		UNCLASSIFIED	264488, 56182575, 22278985, 35696288, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 29331827, 35696052, 29331828, 29148498, 29331830, 265006, 265007, 265008, 60432228, 33657402, 55812038, 87168474, 265010, 265011, 265017, 265018, 265019, 264605, 264681, 264288, 264369, 52844229, 21908765, 21908768, 21906767, 21906768, 21906769, 265020, 265022, 264691, 264692, 33657109, 18108370, 18108374, 55810764, 35695855, 264634, 60431850, 264639, 56182323, 18108382, 18108385, 65274727, 22279002, 264584
2741	88047518 (5481, 5482)	Novel Protein sim. GBank gij3242764 (AC005154) - similar to protein U28928 (PID:g861308) [Homo sapiens]		UNCLASSIFIED	22278988, 52844045, 52644228, 21906768, 21906769, 265020, 60170815, 264691
2742	87648644 (5483, 5484)	Novel Protein sim. GBank gij4758412 [refjNP_004472.1] pGALN - UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylglucosaminyltransferase 2 [Gallinac-T2]	Contains protein domain (PF00852) - Similarity to lectin domain of ricin beta-chain, 3 copies.	transferase	264259, 264805, 264758, 55812038, 264369, 29148627
2743	87827891 (5485, 5486)	Novel Protein sim. GBank gij4468311 [embjCAB37992] - (AL031432) dJ465N24.2.1 (PUTATIVE novel protein) (isoform 1) [Homo sapiens]		UNCLASSIFIED	35696288, 264259, 264906, 264908, 265006, 60433438, 265017, 18108351, 264448, 264784, 264288, 21908765, 21908767, 264690, 264691, 264693, 263969, 263971, 35695855, 264637, 264558, 18108382, 60432113
2744	94126030 (5487, 5488)	Novel Protein sim. GBank gij3880433 [embjCAA91399] - (Z66521) similar to mitochondrial RNA splicing MSRA like protein; cDNA EST EMBL: C09217 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	kinase	18108374, 264488, 56182435, 21908765, 35696423, 35695917, 35695855, 265020, 265021, 264511, 265008, 264490, 264556, 264259, 264557, 56182323, 264558, 264559, 18108383, 29331824, 18108385, 33657109, 29331826, 21908764, 29331827, 29331828, 33657349, 87168516, 265018, 264905, 264482, 264448, 264488, 264369, 264288

2745	87740125 (5489, 5490)	Novel Protein sim. GBank gij4405785[gijAAD19826] - (AF038963) RNA helicase [Homo sapiens]	Contains protein domain (PF00271) - helicase domain Helicases conserved C-terminal domain	35696286, 284509, 284905, 284907, 284908, 284909, 284510, 284512, 265008, 284758, 284801, 285017, 284604, 284783, 284288, 284886, 284789, 284693, 35698423, 35695855, 284634, 284638, 284563, 284564, 284565
2746	85418601 (5491, 5492)	Novel Protein sim. GBank gij4768738[ref]NP_004680.1pMTA1 - metastasis associated 1	Contains protein domain (PF00320) - GATA zinc finger	22278996, 22278998, 22278999, 29331822, 29331826, 29331827, 35696052, 29331828, 284905, 284908, 284907, 284908, 284909, 52844045, 285008, 60170831, 284598, 55812038, 265018, 284683, 284288, 21908765, 21908787, 21908768, 21908769, 265020, 284690, 33657023, 284693, 33657109, 18108388, 18108374, 284558, 18108385, 22278000, 284563
2747	94112677 (5493, 5494)	Novel Protein sim. GBank gij4557803[ref]NP_000282.1pNPC1 - Niemann-Pick disease, type C1	glycoprotein	284589, 52844507, 18108394, 22278995, 35696286, 22278997, 22278999, 52845080, 29331824, 56182181, 29331826, 29331827, 35696052, 284907, 284908, 284908, 265009, 33109954, 55811386, 87168474, 285010, 87168559, 284603, 265019, 284760, 284686, 284768, 21908769, 35695917, 60170615, 284692, 33657023, 52845129, 27486284, 60431528, 18108374, 35698423, 35695855, 284558, 58182323, 18108385, 284482
2748	91214883 (5495, 5496)	Novel Protein sim. GBank gij4191272[emb]CAA089841 - (AJ012295) spaG protein [Rhizobium etli]	Contains protein domain (PF00646) - F-box domain.	65274572, 29331828, 284112, 284511, 265019, 284760, 284767, 284768, 284769, 21908768, 21908769, 265020, 27486282, 56528488, 87188518, 22278000
2749	87346307 (5497, 5498)			284259, 284908, 284510, 265008, 265009, 284780, 284389, 284768, 284563
2750	87336344 (5499, 5500)	Novel Protein sim. GBank gij1872498 (U74297) - PIUS [Oryctolagus cuniculus]	UNCLASSIFIED	284488, 52844507, 18108396, 58994075, 264259, 29331825, 29331826, 29331827, 29331828, 284508, 285009, 284910, 284591, 284595, 33657084, 265011, 265019, 18108351, 284288, 284686, 284768, 284689, 55811957, 284693, 27486284, 18108370, 18108374, 284558, 18108385, 284482, 284583
2751	87057465 (5501, 5502)		UNCLASSIFIED	29331822, 29331824, 265017, 33657023
2752	88082875 (5503, 5504)	Novel Protein sim. GBank gij3041859 (AC004534) - OG-2 homeodomain protein-like, similar to U65087 (PID:g1575526) [Homo sapiens]	Contains protein domain (PF00046) - Homeobox domain	

2753	84138972 (5505, 5506)	Novel Protein sim. GBank gij3851648 (AF098301) - neural F-box protein NFB42 [Rattus norvegicus]	Contains protein domain (PF00646) - F-box domain.	UNCLASSIFIED	56182575, 56994075, 22278998, 28331822, 29331824, 29331825, 29331826, 285007, 284593, 55812038, 3310954, 18108351, 284288, 58181582, 21908767, 21908768, 285021, 284693, 18108374, 65274791, 284632, 58182323, 22279002, 284563, 284597
2754	84115513 (5507, 5508)	Novel Protein sim. GBank gij535428 (U13738) - calmodulin-like protein [Pisum sativum]	Contains protein domain (PF00038) - EF hand	struct	22278999, 66714117, 29331827, 35696032, 28331828, 284908, 284908, 284909, 265011, 265017, 265018, 265019, 264288, 21908765, 21908767, 265022, 33657023, 284893, 58182323, 18108382, 22279000
2755	88001472 (5509, 5510)	Novel Protein sim. GBank gij2898653 (AC004510) - R30385_2 [Homo sapiens]	Contains protein domain (PF00098) - Zinc finger, C2H2 type	transcriptfactor	
2756	11465808 (5511, 5512)			UNCLASSIFIED	284594
2757	93381580 (5513, 5514)	Novel Protein sim. GBank gij1173539 (U30473) - putative src-like adapter protein; non-catalytic src-like adapter protein containing SH3 and SH2 domains; homolog of mouse SLAP; Method: conceptual translation supplied by author [Homo sapiens]	Contains protein domain (PF00017) - Src homology domain 2	eph	85658542, 284693
2758	78637846 (5515, 5516)	Novel Protein sim. GBank gij3608372 (AF053768) - brain specific contactin-binding protein CBP90 [Rattus norvegicus]		struct	284693
2759	91005312 (5517, 5518)	Novel Protein sim. GBank gij2072200 (U94863) - p40 [Borna disease virus]			65274572, 35698286, 66714117, 29331828, 284508, 58182435, 21908754, 55811957, 284629, 284636, 58182323, 22279002
2760	78824798 (5519, 5520)			UNCLASSIFIED	284908
2761	87639597 (5521, 5522)	Novel Protein sim. GBank gij4914573 (emb CAB43685.1) (AL050380) hypothetical protein [Homo sapiens]	Contains protein domain (PF00023) - Ank repeat	nuclease	22278998, 29331822, 29331830, 265010, 285019, 264288, 21908785, 21908768, 21908769, 265020, 58182323, 22279002, 284563
2762	87592899 (5523, 5524)	Novel Protein sim. GBank gij3136150 (AF050182) - PERIOD 3 [Mus musculus]	Contains protein domain (PF00989) - PAS domain	struct	18108394, 22278998, 264906, 284909, 285006, 285007, 284757, 285010, 285011, 285017, 265019, 18108351, 284448, 284683, 284686, 284768, 265020, 265021, 265022, 284691, 18108362, 284693, 18108365, 33657109, 18108368, 18108370, 18108381, 18108382, 18108384, 18108388, 87168518, 284389, 35698423
2763	87539888 (5525, 5526)	Novel Protein sim. GBank gij3511122 (AF060503) - zinc finger protein [Homo sapiens]	Contains protein domain (PF01352) - KRAB box	transcriptfactor	
2764	84305140 (5527, 5528)	Novel Protein sim. GBank gij2805843 (AF045244) - ribitol kinase [Klebsiella pneumoniae]	Contains protein domain (PF00370) - FGGY family of carbohydrate kinases		52848842, 284259, 29331822, 29331825, 29331826, 29331828, 33656970, 284905, 284907, 29331830, 265006, 265009, 21908754, 265019, 284448, 21908769, 27488282, 58182323, 58528486, 87168518, 284487

2765	94315105 (5528, 5530)	Novel Protein sim. GBank gij4688672[emb]CAA17888.2] - (AL022018) /prediction=(method: /prediction=(method: /match=(desc: [Drosophila melanogaster]		264488, 65274572, 22278995, 35696286, 22278998, 22278999, 264259, 29331822, 29331824, 35696052, 264509, 264509, 264905, 264906, 264907, 264908, 264909, 264512, 264910, 265009, 264592, 264595, 264758, 55812038, 33109954, 265010, 87168559, 264800, 265018, 264760, 264761, 264762, 264763, 264448, 264764, 264288, 264766, 264767, 264768, 58181582, 21908784, 21908785, 21908788, 35695917, 285021, 284691, 284692, 33857023, 33657109, 284628, 18108370, 284629, 18108374, 55811576, 35696423, 35695855, 264631, 264634, 264635, 264636, 264637, 264638, 264558, 264639, 83373044, 87168518, 22278000, 22279002, 264563, 264482
2766	94315109 (5531, 5532)	Novel Protein sim. GBank gij5441611[emb]CAB48854.1] - (AJ388555) hypothetical protein [Canis familiaris]	Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	264488, 65274572, 60432289, 264907, 264909, 264511, 264512, 60433358, 264288, 264685, 264689, 35695917, 265022, 264693, 264628, 65274791, 264635, 264555, 264556, 264557, 264636, 264558, 264559, 83373044, 60432113
2767	80204297 (5533, 5534)	Novel Protein sim. GBank gij1079451[pr]AJ55463 - tropomodulin, skeletal muscle - chicken	struct	264112, 263974, 264558
2768	94322238 (5535, 5536)	Novel Protein sim. GBank gij5441322[emb]CAB46721.1] - (AL031427) dJ167A19.1 (novel protein) [Homo sapiens]	UNCLASSIFIED	264488, 263994, 264489, 35696286, 22278998, 22278999, 264259, 29331822, 60432289, 29331828, 35696052, 264508, 264509, 264905, 264906, 264907, 29331830, 264908, 264909, 264510, 264511, 264512, 265008, 264910, 265009, 264591, 264592, 264593, 33857402, 264594, 264595, 264757, 264596, 265011, 264604, 265019, 264760, 264681, 18108351, 264764, 264288, 264766, 264686, 264768, 18108357, 264769, 264689, 21908768, 21908769, 35695917, 265021, 265022, 33657023, 33657109, 264628, 264629, 18108374, 35696423, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264555, 264638, 264639, 83373044, 264583, 264584, 264585, 264566, 264488, 264567

2769	95311088 (5537, 5538)	Novel Protein sim. GBank gjl5419859[emb]CAB46375.1] - (AL088725) hypothetical protein [Homo sapiens]		tubulin	264488, 56182575, 264259, 29331822, 29331824, 66714117, 29331825, 60432289, 29331828, 264908, 29331830, 56182435, 264592, 33657402, 264448, 264389, 264288, 60170615, 264891, 33657023, 264682, 33657109, 18108374, 55811576, 264634, 264636, 56182323, 83373044, 60432113, 18108398, 22278995, 22278998, 22278999, 264105, 265006, 265019, 18108351, 264687, 21908765, 18108364, 264629, 18108374, 264631, 18108385, 18108388
2770	87730182 (5539, 5540)	Novel Protein sim. GBank gjl5701965[emb]CAB52157.1] - (AL108738) WD repeat protein [Schizosaccharomyces pombe]	Contains protein domain (PF00400) - WD domain, G-beta repeat	kinase	18108398, 22278995, 22278998, 22278999, 264105, 265006, 265019, 18108351, 264687, 21908765, 18108364, 264629, 18108374, 264631, 18108385, 18108388
2771	88084071 (5541, 5542)	Novel Protein sim. GBank gjl3093433 (AC004125) - Unknown gene product [Homo sapiens]		UNCLASSIFIED	264259, 29331822, 29331824, 29331825, 264369
2772	95357309 (5543, 5544)	Novel Protein sim. GBank gjl4885531[re]NP_005465.1[pNY C - histone deacetylase 5	Contains protein domain (PF00850) - Histone deacetylase family	histone	264259, 29331822, 29331824, 66714117, 60432289, 29331827, 264805, 265009, 264592, 55812038, 65274444, 264786, 21908769, 33657109, 263978, 264555, 264638, 264557, 83373044, 264593, 264564, 264486
2773	84138994 (5545, 5546)	Novel Protein sim. GBank gjl3288888 (AC005253) - R26445_1 [Homo sapiens]		UNCLASSIFIED	18108374, 264686, 264687, 263976, 56182435, 264689, 55810784, 21908786, 35686423, 55811576, 65274791, 56181686, 55811957, 35695855, 264110, 265021, 264112, 265022, 265008, 265008, 264092, 264084, 60431850, 264637, 264691, 264259, 264692, 263981, 264594, 60433356, 56182323, 264693, 264757, 56182181, 55812038, 29331825, 60424268, 18108385, 29331826, 29331827, 27486261, 29331828, 35686052, 55811386, 264107, 60432113, 265017, 55811150, 18108351, 264681, 264606, 18108370, 264484, 264682, 20281068, 264448, 68712502, 264683, 264764, 264288, 264684, 264768, 263974
2774	87819908 (5547, 5548)	Novel Protein sim. GBank gjl465852[sp]P34388[YL3 CAEL - HYPOTHETICAL 70.7 KD PROTEIN F09G8.3 IN CHROMOSOME III	Contains protein domain (PF00380) - Ribosomal protein S9/S16	ribosomalprot	22278995, 35686286, 22278998, 22278999, 264259, 29331828, 60432289, 35686052, 264112, 33657402, 21908754, 87168559, 265017, 265018, 264288, 21908766, 21908767, 21908769, 35685917, 265020, 265021, 33657023, 33657109, 18108370, 263978, 35686423, 35695855, 87168518, 22278000, 264482

2775	95307987 (5549, 5550)	Novel Protein sim. GBank gi 4689132 gb AAD27775.1 AF077042) 30S ribosomal protein S7 homolog [Homo sapiens]	Contains protein domain (PF00177) - Ribosomal protein S7p/S5e	ribosomal prot	264488, 22278995, 56894075, 22278898, 35695286, 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 60432289, 29331827, 29331828, 35696052, 265007, 21906754, 265017, 265019, 284448, 264682, 264369, 264288, 18108354, 52644228, 264689, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265021, 265022, 60170815, 264691, 18108370, 35696423, 65274791, 35695855, 264634, 60431850, 60170394, 56182323, 264558, 18108388, 22279000, 264563, 264565, 264486, 264567
2776	87701587 (5551, 5552)				56182575, 22278998, 22278999, 264259, 29331822, 29331824, 264608, 29331830, 264510, 33657402, 21906754, 55811388, 265017, 265019, 264448, 264288, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 265021, 265022, 60170615, 55810764, 55811576, 264555, 56526486, 22279000
2777	79816729 (5553, 5554)			UNCLASSIFIED	264907, 264766
2778	8212411 (5555, 5556)			UNCLASSIFIED	264907, 264593, 264760, 264628
2779	87649729 (5557, 5558)	Novel Protein sim. GBank gi 4680711 gb AAD27745.1 AF132970) CGI-36 protein [Homo sapiens]		UNCLASSIFIED	22278997, 264259, 29331824, 66714117, 35696052, 265008, 264512, 264448, 264288, 29148627, 18108364, 20281149, 18108370, 264629
2780	94678397 (5559, 5560)	Novel Protein sim. GBank gi 4758524 ref NP_004825.1 pHGX - HPK/GCK-like kinase	Contains protein domain (PF00780) - CNH domain	kinase	29146499, 65274791, 264634, 264639
2781	91220037 (5561, 5562)	Novel Protein sim. GBank gi 469332 gb AAD21222 - (AF068502) ubiquitin specific protease UBP43 [Mus musculus]	Contains protein domain (PF00443) - Ubiquitin carboxyl-terminal hydrolase family 2	ubiquitin	60424179, 29331824, 60424269, 66714117, 29331828, 56182435, 87168474, 265017, 264764, 56181562, 21906765, 21906766, 21906768, 35695917, 265020, 33657023, 35695855, 56182323, 87168518
2782	94233146 (5563, 5564)	Novel Protein sim. GBank gi 4505013 ref NP_002310.1 pLRN - leucine-rich neuronal protein	Contains protein domain (PF00560) - Leucine Rich Repeat	struct	65274572, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 29331828, 60432289, 29331828, 264905, 264907, 264908, 264909, 52644045, 265009, 265017, 265018, 264604, 265019, 264760, 264683, 264288, 264766, 264685, 264686, 284768, 52644229, 264689, 21906768, 265020, 265021, 264691, 18108362, 264692, 33657023, 264693, 33657109, 33657349, 18108370, 264628, 263978, 35695855, 264557, 56182323, 83373044, 18108385

2783	80016629 (5565, 5568)	Novel Protein sim. GBank gij28831[sp]P39188[ALU1_HUMAN - III] ALU SUBFAMILY J WARNING ENTRY III		lm7	264909, 264628, 263978, 263981
2784	87014360 (5567, 5568)				264259, 28331822, 28331824, 28331825, 264482
2785	88071830 (5569, 5570)	Novel Protein sim. GBank gij2134933[prj]S58890 - collapsin response mediator protein - human		UNCLASSIFIED	22278999, 264908, 264758, 265018, 264769, 21908765, 21908768, 21908769, 265020, 284584
2786	87408542 (5571, 5572)	Novel Protein sim. GBank gij2073564 (U80223) - eukaryotic initiation factor eIF-2 alpha kinase; DGCN2 [Drosophila melanogaster]	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	kinase	284805, 265017
2787	87901268 (5573, 5574)	Novel Protein sim. GBank gij5174507[ref]NP_006020.1[pMA1] - paraneoplastic neuronal antigen		UNCLASSIFIED	264768, 21908765, 21908768, 55811957, 22278999, 264093, 60170815, 264259, 29331822, 18108385, 29331824, 33657109, 29331827, 35696052, 264100, 264105, 264908, 263977, 55811576, 264639, 264637, 60433438, 60432113, 265017, 265019, 22279002, 55811150, 264369, 264288
2788	88090844 (5575, 5576)	Novel Protein sim. GBank gij3252826 (AC004382) - Unknown gene product [Homo sapiens]		UNCLASSIFIED	
2789	83491275 (5577, 5578)	Novel Protein sim. GBank gij2495729[sp]Q82556[Y281_HUMAN - HYPOTHETICAL PROTEIN KIAA0281 (HA8725)]		UNCLASSIFIED	264103, 21908769, 264693
2790	87602784 (5579, 5580)	Novel Protein sim. GBank gij5101772[emb]CAB45135.1] - (AJ242978) p821 [Homo sapiens]		struct	264488, 264769, 264910, 264631, 264638, 264980, 264691, 264259, 264638, 29331824, 264508, 264509, 264905, 264583, 264908, 264628, 18108370, 264907, 264764, 264908, 264288, 264909
2791	88083195 (5581, 5582)	Novel Protein sim. GBank gij2911266 (AC002550) - Unknown gene product [Homo sapiens]			21806764, 18108368
2792	95083783 (5583, 5584)	Novel Protein sim. GBank gij2854163[gb]AAC02581.1] - (AF045842) No definition line found [Caenorhabditis elegans]			22278996, 22278997, 264259, 29331822, 29331824, 29331825, 66714117, 29331826, 60432289, 29331827, 35698052, 264908, 66712502, 29331830, 264909, 60432229, 60433356, 60433438, 33109954, 265010, 265011, 265017, 265018, 265019, 264682, 264448, 264288, 264369, 264768, 52844229, 21908765, 21908768, 265020, 285021, 33657023, 263974, 18108374, 65274791, 35895855, 264636, 264558, 264558, 56182323, 83373044, 18108385, 56526486, 60432113, 22279000, 264587
2793	87425476 (5585, 5586)			UNCLASSIFIED	264259, 60432289, 66712502, 265009, 264638
2794	85794830 (5587, 5588)			UNCLASSIFIED	264689, 265022, 264691, 18108388, 264567

2795	9533-888 (5589, 5590)	Novel Protein sim. GBank gi 5454146 ref NP_008348.1 pUBE2 - ubiquitin-conjugating enzyme E2E 3 (homologous to yeast UBC4/5)	Contains protein domain (PF00170) - Ubiquitin-conjugating enzyme	ubiquitin	65274572, 56182575, 35698286, 22278988, 22278999, 60432049, 264259, 29331822, 29331825, 66714117, 60432289, 29331828, 29331827, 35698052, 29331828, 264907, 66712502, 56182435, 264511, 265007, 264512, 264910, 60432229, 60433358, 60433438, 33108954, 8568542, 265018, 265019, 264288, 264688, 21906784, 21906785, 21906788, 21906788, 21906789, 55811957, 265020, 265021, 265022, 52644150, 33657023, 264693, 85274620, 33657109, 35696423, 55811576, 85274781, 56182323, 56526486, 60432113, 22278002, 264482, 264563, 264484, 264567
2796	94648857 (5591, 5592)	Novel Protein sim. GBank gi 4680831 gb AAD27715.1 AF13294 - (AF13294) CGI-08 protein [Homo sapiens]		UNCLASSIFIED	18108394, 65274572, 56182575, 56994075, 22278999, 264480, 60432049, 264259, 29331822, 29331824, 29331826, 35696052, 264509, 264908, 264907, 264908, 66712502, 56182435, 264510, 265006, 264512, 265007, 265008, 264910, 265009, 264591, 264592, 60433358, 264594, 264595, 55812038, 264596, 21906754, 60174639, 87188474, 265010, 265011, 265017, 265018, 265019, 55811150, 264782, 18108351, 264448, 264682, 264369, 264288, 264685, 264766, 264687, 56181562, 264769, 264689, 21906765, 21906768, 21906767, 21906768, 21906769, 55811957, 35695917, 265020, 265021, 60170815, 52644150, 264692, 33657023, 18108362, 264693, 65274620, 33657109, 33657182, 27466265, 33657349, 18108374, 35696423, 65274791, 35695955, 264556, 264557, 56182323, 264558, 60170394, 83373044, 65274727, 87168518, 22278000, 264563, 264564, 264565, 264566, 264567
2797	95110790 (5593, 5594)	Novel Protein sim. GBank gi 4838557 gb AAD31040.1 - (AF143858) potassium channel modulatory factor DEBT-91 [Mus musculus]	Contains protein domain (PF00569) - Zinc finger present in dystrophin, CBP/p300	UNCLASSIFIED	56182375, 22278995, 22278996, 22278997, 29331827, 28146489, 264509, 264908, 56182435, 264757, 21906754, 265010, 265017, 265019, 264681, 264682, 264683, 264686, 21906765, 21906767, 21906768, 21906769, 28146829, 265020, 52644150, 264690, 33657182, 264629, 18108376, 56182323, 22278002, 264563
2798	86198005 (5595, 5596)	Novel Protein sim. GBank gi 2852645 (AF007160) - unknown [Homo sapiens]		UNCLASSIFIED	264486, 264480, 264259, 264448, 20281149, 20281152, 264558, 264557, 264558, 264559, 264483, 264486, 264567

2799	88080651 (5597, 5598)	Novel Protein sim. GBank gij3252825 (AC004382) - Unknown gene product [Homo sapiens]		UNCLASSIFIED	22278996, 22278997, 22278998, 22278999, 60432049, 264259, 28331828, 35698052, 29331828, 264909, 60433356, 33657402, 33109954, 87168474, 264448, 52644229, 21908768, 21908767, 21908768, 35895917, 265020, 265021, 265022, 52644150, 35695855, 264634, 60432113, 22279000, 284488, 56994075, 264259, 20281089, 29331825, 29331827, 264905, 56182435, 265006, 265011, 87168559, 265017, 265019, 264448, 264288, 264786, 264888, 60170815, 284691, 264692, 27486265, 264828, 264629, 264636, 264557, 264558, 264559, 87168518, 264584, 264568, 264567, 265007, 264687
2800	88316461 (5599, 5600)	Novel Protein sim. GBank gij4240301[dbj]BAA74929.1] - (AB020713) KIAA0906 protein [Homo sapiens]		glycoprotein	
2801	86068814 (5601, 5602)		Contains protein domain (PF00627) - UBA domain		
2802	88082477 (5603, 5604)	Novel Protein sim. GBank gij2337865 (AC002464) - organic cation transporter; 50% similarity to JC4884 (PID:g2143692) [Homo sapiens]	Contains protein domain (PF00083) - Sugar (and other) transporter	transport	264448, 35695855
2803	78577448 (5605, 5606)	Novel Protein sim. GBank gij4559368[gib]AAD23029.1[AC006585] - (AC006585) hypothetical protein [Arabidopsis thaliana]	Contains protein domain (PF01585) - G-patch domain	UNCLASSIFIED	284639
2804	57111131 (5607, 5608)			peptidase	264568
2805	87398488 (5609, 5610)				284092, 264259, 29331822, 29331824, 29331827, 29331828, 264508, 265007, 265009, 21906754, 264369, 264288, 264632, 60170394, 264563, 264482
2806	87898951 (5611, 5612)	Novel Protein sim. GBank gij1168973[sp]P44403[CLPB_HAEIN - CLPB PROTEIN]		UNCLASSIFIED	22278995, 22278996, 22278998, 264259, 29331822, 29331824, 29331826, 29331827, 29146498, 285008, 265009, 80433438, 265017, 265018, 265019, 264448, 264288, 21908765, 21908767, 21908769, 29148629, 265022, 52644150, 56182323

2807	91720702 (5813, 5614)	Novel Protein sim. GBank gl 4468310 emb CAB379811 - (AL031432) dJ46SN24.1 (PUTATIVE novel protein similar to predicted yeast and worm proteins) [Homo sapiens]		UNCLASSIFIED	52644507, 52645156, 52646842, 18108398, 56182575, 22278994, 22278995, 22278996, 56994075, 35698286, 22278997, 22278998, 22278999, 264259, 32845080, 29331822, 29331824, 29331825, 29331826, 35696052, 29331828, 33656970, 264100, 264105, 284907, 52644045, 60433356, 284594, 60433438, 52646317, 21908754, 33109954, 33657084, 52644296, 87188474, 265011, 87188559, 265017, 265018, 265019, 18108351, 264448, 264763, 264887, 52644229, 21908765, 21908766, 21908767, 21908768, 21908769, 35695917, 265021, 265022, 52644150, 33657023, 52645128, 33657109, 33657182, 27486261, 27486262, 27486264, 27486265, 35695763, 18108370, 18108372, 18108374, 18108376, 263977, 18108377, 35696423, 35695855, 52644332, 83373044, 18108385, 18108387, 87188518, 60432113
2808	95359111 (5615, 5616)	Novel Protein sim. GBank gl 5541883 emb CAB51071.1 - (AL098857) hypothetical protein [Homo sapiens]	MHC		60432289, 264510, 265010, 265017, 265018, 264681, 264686, 265021, 264690, 22279000, 264586
2809	88083530 (5617, 5618)	Novel Protein sim. GBank gl 272561 (AC004002) - similar to ciliary dynein beta heavy chain; 76% Similarity to P23098 (PID:g118965) [Homo sapiens]	ATPase associated		18108351
2810	87259032 (5619, 5620)			UNCLASSIFIED	284569, 22278996, 264091, 264259, 29331828, 29148499, 29148629, 29148784, 264693
2811	91235945 (5621, 5622)	Novel Protein sim. GBank gl 3264583 (AC005189) - match to ESTs H97758 (NID:g118843) and AA085546 (NID:g1828773) [Homo sapiens]		UNCLASSIFIED	22278999, 264259, 66712502, 264693
2812	88093334 (5623, 5624)			UNCLASSIFIED	264106
2813	91218755 (5625, 5626)	Novel Protein sim. GBank gl 4240273 db BAA74915.1 - (AB020699) KIAA0892 protein [Homo sapiens]		UNCLASSIFIED	56182575, 29331828, 264906, 66712502, 55811386, 265017, 265018, 265019, 264683, 265020, 87188518, 60432113
2814	90880906 (5627, 5628)	Novel Protein sim. GBank gl 3548791 (AC005620) - R33580.1 [Homo sapiens]	transcription factor		65274572
2815	79774521 (5629, 5630)			UNCLASSIFIED	264907, 264909
2816	95358229 (5631, 5632)	Novel Protein sim. GBank gl 5420389 emb CAB46880.1 - (AJ243460) proteophosphoglycan [Leishmania major]		UNCLASSIFIED	284488, 35696286, 29331825, 29331826, 284508, 264509, 264905, 264906, 264907, 264908, 264909, 264511, 264910, 265011, 264760, 264681, 264766, 264769, 264689, 21908765, 264693, 264628, 18108370, 264629, 264631, 264634, 264563, 264564, 284566, 264488

2817	87748542 (5633, 5634)	Novel Protein sim. GBank gi1293846 (U56966) - coded for by <i>C. elegans</i> cDNA yk30b3.5; coded for by <i>C. elegans</i> cDNA yk30b3.3 [<i>Caenorhabditis elegans</i>]			264259, 29331822, 29331827, 264508, 264509, 284905, 264907, 264908, 264909, 58182435, 264510, 265008, 264511, 264512, 284593, 264756, 265010, 264760, 264761, 264764, 264288, 264687, 264769, 55811957, 35895917, 33657109, 263978, 264634, 264636, 264639, 264564, 264585, 264566, 264488, 264587
2818	88073579 (5635, 5636)	Novel Protein sim. GBank gi1549986 (U13149) - possible apospory-associated protein [<i>Pennisetum ciliare</i>]	Contains protein domain (PF00023) - Ank repeat	UNCLASSIFIED	66712502
2819	87703527 (5637, 5638)	Novel Protein sim. GBank gi14929773 [gb AAD34147.1 AF15209] CGI-05 protein [<i>Homo sapiens</i>]	Contains protein domain (PF00919) - Uncharacterized protein family UPF0004	UNCLASSIFIED	264909, 264511
2820	87765744 (5639, 5640)	Novel Protein sim. GBank gi14929773 [gb AAD34147.1 AF15209] CGI-05 protein [<i>Homo sapiens</i>]	Contains protein domain (PF00919) - Uncharacterized protein family UPF0004	UNCLASSIFIED	18108394, 52846365, 52844045, 264112, 265009, 21906754, 265017, 18108351, 264683, 264389, 264688, 264769, 21906769, 33657023, 33657109, 18108370, 18108374, 18108377, 18108385
2821	95320511 (5841, 5842)	Novel Protein sim. GBank gi1398144 [sp P02747 C1QC_HUMAN - COMPLEMENT C1Q SUBCOMPONENT, C CHAIN PRECURSOR	Contains protein domain (PF00386) - C1q domain	complement	264638
2822	94260221 (5843, 5844)	Novel Protein sim. GBank gi12224671 [db BAA20820] - (AB002363) KIAA0385 [<i>Homo sapiens</i>]		UNCLASSIFIED	264259, 29331822, 60432289, 264905, 60433356, 33657402, 265011, 264760, 21908765, 60170615, 264638
2823	95320513 (5845, 5846)	Novel Protein sim. GBank gi1398144 [sp P02747 C1QC_HUMAN - COMPLEMENT C1Q SUBCOMPONENT, C CHAIN PRECURSOR	Contains protein domain (PF00386) - C1q domain	complement	264766
2824	95320515 (5847, 5848)	Novel Protein sim. GBank gi1398144 [sp P02747 C1QC_HUMAN - COMPLEMENT C1Q SUBCOMPONENT, C CHAIN PRECURSOR		UNCLASSIFIED	264907
2825	18742170 (5849, 5850)	Novel Protein sim. GBank gi13859883 [emb CAA22020] - (AL033503) conserved hypothetical protein [<i>Candida albicans</i>]			264760
2826	94311805 (5851, 5852)	Novel Protein sim. GBank gi13859883 [emb CAA22020] - (AL033503) conserved hypothetical protein [<i>Candida albicans</i>]			52645156, 22278994, 35696286, 22278987, 22278998, 52645080, 29331822, 29331824, 29331825, 29331827, 35696052, 29331828, 33656970, 52644045, 52646317, 33657084, 52644286, 265017, 265018, 265019, 264288, 21906764, 21906765, 21906768, 21906769, 21906769, 35695917, 52644150, 33657109, 33657182, 27486261, 27486262, 33657349, 27486265, 35695763, 35696423, 35695855, 52644332, 18108385, 87168518, 264484

2827	95320519 (5653, 5654)	Novel Protein sim. GBank gi 398144 sp P02747 C1QC_HUMAN - COMPLEMENT C1Q SUBCOMPONENT, C CHAIN PRECURSOR	Contains protein domain (PF00386) - C1q domain	complement	264488, 263994, 264489, 65274572, 29331822, 66714117, 29331827, 29331828, 264508, 264905, 264509, 264906, 264907, 264908, 264909, 264510, 265008, 264511, 265007, 264512, 265008, 264910, 264591, 60432229, 264592, 264593, 264595, 264596, 21908754, 265011, 264600, 264601, 264602, 265017, 264604, 264605, 264288, 264766, 264767, 264689, 55811957, 264534, 264690, 264891, 264892, 264693, 263972, 264629, 35695855, 264631, 264634, 264635, 264555, 264638, 264637, 264556, 264638, 264556, 83373044, 18108385, 60432113, 22279002, 264563, 264564, 264565, 264566, 264567, 264486, 18108391
2828	91228615 (5655, 5656)	Novel Protein sim. GBank gi 3598974 (AF077000) - protein tyrosine phosphatase TD14 [Rattus norvegicus]	Contains protein domain (PF00102) - Protein-tyrosine phosphatase	phosphatase	29331822, 35696052, 264104, 264908, 265007, 264591, 265010, 265011, 265019, 264768, 264686, 55811957, 18108370, 18108374, 55810764, 35696423, 55811576, 56182323, 83373044, 87168518
2829	87651244 (5657, 5658)	Novel Protein sim. GBank gi 4880889 gb AAD27734.1 AF13295 - (AF132959) CGI-25 protein [Homo sapiens]			22278996, 22278997, 264091, 264093, 60432049, 264259, 29331822, 29331825, 29331827, 29331828, 264905, 264509, 66712502, 264510, 264511, 264593, 60433438, 21908754, 265011, 264603, 18108351, 264288, 21908765, 21908768, 21908769, 2648629, 52644150, 264693, 33657109, 18108374, 264634, 18108385, 60432113, 22279000, 264585, 264488 265008, 265018, 264639, 22279002
2830	88087109 (5659, 5660)	Novel Protein sim. GBank gi 2498667 sp Q61200 NPH1_MOUSE - NEUREXOPHILIN 1			
2831	87614717 (5661, 5662)			UNCLASSIFIED	265017
2832	87631809 (5663, 5664)			UNCLASSIFIED	22278997, 22278999, 52646317, 264288, 264688, 21908767, 60431528, 264638, 22279000
2833	87612938 (5665, 5666)	Novel Protein sim. GBank gi 5262615 emb CAB45747.1 - (AL080158) hypothetical protein [Homo sapiens]		UNCLASSIFIED	264555, 264556, 264558
2834	86974703 (5667, 5668)	Novel Protein sim. GBank gi 2224567 dbj BAA20772 - (AB002311) KIAA0313 [Homo sapiens]			263972
2835	87715712 (5669, 5670)	Novel Protein sim. GBank gi 4589532 dbj BAA76788.1 - (AB023161) KIAA0944 protein [Homo sapiens]		ATPase associated	60432289, 29331828, 265008, 265010, 265017, 264448, 55811957, 265020, 18108370
2836	85724748 (5671, 5672)	Novel Protein sim. GBank gi 2351568 (U76618) - N-RAP [Mus musculus]	Contains protein domain (PF00412) - LIM domain containing proteins	transcript factor	264259, 264112, 265010, 264762, 264784, 293974, 264555, 264558, 264559

2837	87766482 (5673, 5674)	Novel Protein sim. GBank gij5420387[embjCAB46678.1] - (AJ243459) proteophosphoglycan [Leishmania major]	UNCLASSIFIED	18108394, 22278997, 22278998, 29331822, 284907, 284909, 285006, 265007, 265018, 265019, 264682, 264288, 21906766, 21906767, 55811957, 35695917, 18108374, 56182323, 22279000, 22279002
2838	87775392 (5675, 5676)	Novel Protein sim. GBank gij973378 (U31263) - core protein [Hepatitis C virus]	UNCLASSIFIED	18108394, 18108397, 264259, 29331828, 285007, 265019, 264448, 18108368, 20281149, 264565, 264567
2839	85799317 (5677, 5678)	Novel Protein sim. GBank gij1575515 (U64899) - thrombospondin-related anonymous protein [Plasmodium gallinaceum]	UNCLASSIFIED	284555
2840	87774665 (5679, 5680)	Novel Protein sim. GBank gij2224805[dbjBAA20760] - (AB002330) KIAA0332 [Homo sapiens]		284509, 284511, 265011, 264288, 284769, 265020, 264634, 264636, 264556
2841	86982568 (5681, 5682)	Novel Protein sim. GBank gij2224805[dbjBAA20760] - (AB002330) KIAA0332 [Homo sapiens]	dna_rna_bind	56182575, 35696052, 264907, 264908, 264909, 264910, 264593, 264595, 264766, 265022, 264691, 33657182, 35695763, 18108370, 35695855, 264631, 264559, 264563, 264567
2842	80080086 (5683, 5684)	Novel Protein sim. GBank gij578857[embjCAB51350.1] - (AL050306) dJ475B7.1 (novel KIAA0615 and KIAA0323 LIKE protein) [Homo sapiens]	UNCLASSIFIED	264600
2843	91012404 (5685, 5686)	Novel Protein sim. GBank gij578857[embjCAB51350.1] - (AL050306) dJ475B7.1 (novel KIAA0615 and KIAA0323 LIKE protein) [Homo sapiens]		284908, 264907, 264908, 284909, 264910, 284784, 35695855, 83373044, 18108385
2844	56731154 (5687, 5688)	Novel Protein sim. GBank gij585123[spjQ08878]FBLC_MOUSE - FIBULIN-1, ISOFORM C PRECURSOR (BASEMENT-MEMBRANE PROTEIN 90) (BM-90)	Contains protein domain (PF00008) - EGF-like domain	284685
2845	94321719 (5689, 5690)	Novel Protein sim. GBank gij5420387[embjCAB46678.1] - (AJ243459) proteophosphoglycan [Leishmania major]	homeobox	29146498, 87168474, 264686, 35696423, 83373044, 264564
2846	88318613 (5691, 5692)	Novel Protein sim. GBank gij5305263[gbjAAD41995.1]AC00823 - (AC008233) unknown protein [Arabidopsis thaliana]	UNCLASSIFIED	29331830, 284909, 285008, 265011, 87168559, 264629, 264556
2847	81811757 (5693, 5694)	Novel Protein sim. GBank gij3398676 (AC005390) - R31180.1 [Homo sapiens]	UNCLASSIFIED	284908, 264766, 264769, 264629, 264637, 284566
2848	87612943 (5695, 5696)	Novel Protein sim. GBank gij5262615[embjCAB45747.1] - (AL080156) hypothetical protein [Homo sapiens]	UNCLASSIFIED	264490, 264259, 264508, 264805, 264907, 264510, 285007, 265008, 284591, 264592, 284593, 264594, 284595, 55812038, 265010, 265011, 284604, 264763, 264764, 264765, 264766, 264686, 264628, 264629, 264555, 264636, 264556, 264557, 264638, 264558, 284559, 284563, 264566, 284567
2849	88084283 (5697, 5698)	Novel Protein sim. GBank gij3342218 (AC004131) - Unknown gene product [Homo sapiens]		28331822, 35696052, 264509, 264908, 265007, 264594, 265018, 264288, 263972, 35696423, 18108384, 56526486, 18108390, 35696286, 265008, 265009, 265018, 264288, 35695917, 264693, 18108374, 35695855, 22279000
2850	87623636 (5699, 5700)		UNCLASSIFIED	

2851	87820548 (5701, 5702)	Novel Protein sim. GBank gl 4321619 gb AAD15788.1 - (AF051098) seven transmembrane domain orphan receptor [Mus musculus]		UNCLASSIFIED	264906, 264907, 56182435, 264758, 55811386, 285010, 18108351, 284448, 264369, 21908765, 284691, 264692, 284693, 35895855, 284556, 18108385, 284567, 264591
2852	86987023 (5703, 5704)	Novel Protein sim. GBank gl 1825729 (U88308) - similar to drosophila membrane protein PATCHED SP:P18502 (PID:g129645) [Caenorhabditis elegans]		UNCLASSIFIED	
2853	87784630 (5705, 5706)	Novel Protein sim. GBank gl 2702347 (AF027503) - putative membrane-associated guanylate kinase 1 [Mus musculus]	Contains protein domain (PF00397) - WW domain	kinase	56182575, 55811150, 264690, 27486262, 27486265, 284632, 56182323, 56528486, 22279002
2854	88033557 (5707, 5708)	Novel Protein sim. GBank gl 2785825 (AC004021) - kelch protein; ring canal component involved in cytoplasmic bridges; 77% Similarity to A45773 (PID:g1079098) [Homo sapiens]	Contains protein domain (PF01344) - Kelch motif	dna_ma_bind	35696286, 29331824, 29331826, 29331828, 284908, 264768, 264693, 22279002, 264482
2855	94723856 (5709, 5710)	Novel Protein sim. GBank gl 1504040 db BAA13219 - (D86983) similar to D.melanogaster peroxidasin(U11052) [Homo sapiens]	Contains protein domain (PF00660) - Leucine Rich Repeat	glycoprotein	22278994, 29331822, 29331824, 29331825, 284906, 264908, 285009, 33109954, 265018, 265019, 284448, 21908765, 265020, 284690, 27486265, 83373044, 22279000, 22279002, 264482
2856	88093359 (5711, 5712)	Novel Protein sim. GBank gl 3284583 (AC005189) - match to ESTs H97758 (NID:g1118643) and AA085546 (NID:g1628773) [Homo sapiens]			21908766, 22278997, 265022, 29331822, 29331828, 27486282, 285007, 265009, 265017, 264482, 264563, 18108351
2857	95348286 (5713, 5714)	Novel Protein sim. GBank gl 3041855 (AC004537) - similar to tumor suppressor p33ING1; similar to AF044078 (PID:g2829208) [Homo sapiens]	Contains protein domain (PF00628) - PHD-finger	struct	22278995, 35696286, 29331824, 29331825, 35696052, 264103, 264108, 56182435, 21906765, 21908769, 265020, 18108368, 35895763, 22279002, 264563
2858	87434748 (5715, 5716)	Novel Protein sim. GBank gl 462585 sp P35227 ME18_HUMAN - DNA-BINDING PROTEIN MEL-18 (ZINC FINGER PROTEIN 144)	Contains protein domain (PF00097) - Zinc finger, C2HC4 type (RING finger)	dna_ma_bind	264569, 264687, 22278995, 22278996, 22278997, 22278999, 264259, 29331826, 29331827, 29331828, 264509, 284905, 264906, 29331830, 264908, 52844045, 264909, 264511, 264512, 265007, 265008, 264910, 265009, 264593, 60433356, 264595, 284758, 21908754, 265010, 265011, 264604, 285018, 264760, 18108351, 264763, 264682, 264764, 264765, 264288, 264369, 264685, 264766, 264768, 18108357, 264769, 21908766, 21908767, 265021, 264534, 60170615, 264691, 264692, 18108370, 264629, 18108374, 264631, 264638, 263981, 18108381, 284558, 18108385, 22279002, 264584, 264586, 264486, 264567, 60424179, 65274572, 29331828, 264905, 264511, 264758, 265011, 21908767, 21908769, 55811957, 265021, 56182323
2859	90937675 (5717, 5718)	Novel Protein sim. GBank gl 4325320 gb AAD17331.1 - (AF124427) claudin-15 [Mus musculus]		UNCLASSIFIED	

2860	87532598 (5719, 5720)	Novel Protein sim. GBank gi 4469186 emb CAB38414.1 - (AL031588) dJ1163J1.2.1 (novel protein similar to C. elegans B0035.16 and bacterial tRNA (5'-Methylaminomethyl-2-thiouridyate)-Methyltransferases) (isoform 1) [Homo sapiens]		UNCLASSIFIED	284259, 28331822, 29331824, 29331825, 29331828, 29331827, 29331828, 284510, 284511, 33109954, 18108351, 284683, 284765, 284389, 284686, 21908785, 284691, 284692, 284693, 18108388, 22279002, 284482
2861	86898507 (5721, 5722)	Novel Protein sim. GBank gi 3941730 (AF108083) - B54 [Homo sapiens]			284369, 284692
2862	87569585 (5723, 5724)	Novel Protein sim. GBank gi 4505013 ref NP_002310.1 pLRLN - leucine-rich neuronal protein		UNCLASSIFIED	284681, 284638
2863	91220421 (5725, 5726)	Novel Protein sim. GBank gi 3249088 (AC004473) - Contains similarity to goliath protein gb M97204 from D. melanogaster. [Arabidopsis thaliana]	Contains protein domain (PF00096) - struct Zinc finger, C2H2 type		56984075, 35696286, 22278998, 29331822, 29331824, 35696052, 29331828, 284108, 284511, 55812038, 33657084, 5581386, 285018, 285019, 21908765, 21908768, 21908769, 35695917, 265020, 265022, 33657023, 33657109, 33657349, 284629, 18108378, 60431850, 56182323, 18108385, 18108387, 87188518, 22279002
2864	87420030 (5727, 5728)	Novel Protein sim. GBank gi 1078451 p A55463 - Tropomodulin, skeletal muscle - chicken	struct		284259, 284910
2865	95312181 (5729, 5730)	Novel Protein sim. GBank gi 438840 (L19048) - MSA-2 [Plasmodium falciparum]			22278995, 21906764, 284482
2866	95103480 (5731, 5732)	Novel Protein sim. GBank gi 585703 sp Q07066 PMP2_RAT - 22 KD PEROXISOMAL MEMBRANE PROTEIN	glycoprotein		65274572, 22278996, 22278998, 22278999, 284259, 29331824, 29331827, 29331828, 60433438, 21906754, 285018, 284448, 284764, 52844150, 83373044
2867	86908001 (5733, 5734)	Novel Protein sim. GBank gi 4580997 gb AAD24571.1 AF12108 - (AF121081) cAMP Inducible 2 protein [Mus musculus]	UNCLASSIFIED		284488, 284768, 21908788, 22278998, 265022, 284259, 284508, 284905, 284907, 284511, 284910, 284635, 284636, 284637, 265011, 265017, 265018, 285019, 284583, 284088, 284566, 284764, 284369, 284567, 284486, 284288, 284768

2868	95303283 (5735, 5738)	Novel Protein sim. GBank gi 1282868 emb CAA63923 - (X94232) t-Cell activation protein [Homo sapiens]		18108392, 264488, 22278994, 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 35696092, 29331828, 264508, 52644045, 264828, 265008, 265007, 265008, 265009, 264591, 60432229, 264593, 60433356, 264595, 21906754, 265017, 265019, 264682, 264369, 21906768, 265021, 21906768, 21906767, 21906768, 265021, 265022, 264691, 33557182, 18108388, 27486261, 27486262, 27486284, 27488265, 18108370, 18108374, 35698423, 35698555, 264832, 58182323, 87188518, 264404, 22278000, 22279002, 264482, 264563, 264564, 264567, 264487
2869	88094412 (5737, 5738)		UNCLASSIFIED	264369
2870	84404574 (5739, 5740)		UNCLASSIFIED	264905, 264908, 264764, 21906769, 264834
2871	88318621 (5741, 5742)	Novel Protein sim. GBank gi 5306263 gb AAD41995.1 AC008233 - (AC008233) unknown protein [Arabidopsis thaliana]	UNCLASSIFIED	264259, 29331822, 60432289, 29331827, 264907, 265008, 265017, 265018, 264682, 264764, 18108354, 265021, 27488265, 264829, 18108387, 264567
2872	95312197 (5743, 5744)	Novel Protein sim. GBank gi 112205 pir B39086 - proline-rich protein 15 - rat	kinase	263981
2873	88094252 (5745, 5746)		UNCLASSIFIED	264488, 18108374, 264768, 264687, 264688, 264689, 35696423, 35696268, 35695917, 264510, 264511, 265007, 264512, 265008, 264910, 264534, 264634, 264635, 264555, 264592, 264259, 264558, 60433438, 60432289, 35696052, 265011, 264800, 264601, 60432113, 264508, 264583, 264482, 264509, 264905, 264906, 264564, 18108351, 264763, 18108370, 264907, 264566, 264908, 264764, 264288, 264567, 264909, 264488, 264768, 18108391

2874	94313548 (5747, 5748)	Novel Protein sim. GBank gij3212854 (AC004005) - unknown protein [Arabidopsis thaliana]	UNCLASSIFIED	52644507, 52845158, 52846365, 56182575, 22278994, 22278995, 56984075, 35696286, 60432049, 264259, 52645080, 29331822, 56182181, 29331824, 60424269, 29331825, 86714117, 29331826, 29331827, 35696052, 29331828, 33656970, 66712502, 52644045, 265007, 265009, 60433356, 264758, 55812038, 18108348, 52646317, 33109954, 33657084, 265017, 264604, 265018, 265019, 264682, 264369, 264288, 264688, 52644229, 21906768, 21906767, 21906768, 55811957, 265020, 265021, 265022, 52644150, 33657023, 52845129, 18108374, 18108378, 35696423, 56182323, 18108387, 87188518, 60432113, 22279000, 22279002, 264563, 264565
2875	88083726 (5749, 5750)	Novel Protein sim. GBank gij2781386 (AC004010) - similar to Leucine-rich transmembrane proteins: 44% similarity to U42767 (PID:g1736918) [Homo sapiens]	glycoprotein	22278996, 22278997, 22278999, 29331828, 29331828, 29146499, 66712502, 265008, 265017, 18108351, 264683, 264689, 21906767, 18108378, 18108377, 55811576, 60170394, 22279000, 264487
2876	88090854 (5751, 5752)	Novel Protein sim. GBank gij2978530 (AC004449) - R33683_2 [Homo sapiens]	Contains protein domain (PF00167) - IgI	
2877	94747029 (5753, 5754)	Novel Protein sim. GBank gij4704208[embjCAB41646.1] - (AL035419) dJ1100H13.1 (putative novel protein) [Homo sapiens]		52846385, 65274572, 56182575, 22278997, 22278998, 22278999, 60432049, 52645080, 60424269, 60432289, 29331827, 35696052, 29331828, 66712502, 52644045, 56182435, 60433358, 33657402, 33657084, 265019, 55811150, 264448, 264368, 21908766, 21906768, 21906769, 265020, 33657023, 33657109, 33657182, 27486262, 264629, 60431528, 55811576, 52644332, 56182323, 264558, 83373044, 18108385, 56326486, 60432113, 22279000, 22279002, 264563
2878	88095309 (5755, 5756)	Novel Protein sim. GBank gij3878775[embjCAB03087] - (Z81077) predicted using GeneFINDER: Similarity to Yeast protein 8248 (TR:G58753.1) [Caenorhabditis elegans]	UNCLASSIFIED	264488, 264259, 29331822, 29331828, 264905, 264509, 264907, 264908, 264510, 265008, 264511, 264512, 33657402, 264758, 21908754, 18108351, 264681, 264682, 264288, 264684, 264685, 264768, 264689, 21908769, 264690, 33657023, 264693, 18108364, 33657109, 264629, 18108374, 264630, 264632, 264558, 264637, 264639, 264558, 18108385, 18108387, 264563, 264564, 264565, 264568, 264488, 264567

2879	87869122 (6757, 5768)	Novel Protein sim. GBank gl 4895145 gb AAD32752.1 - (AF127374) unknown [Streptomyces lavendulae]	Contains protein domain (PF00315) - Uracil-DNA glycosylase	UNCLASSIFIED	18108359, 264259, 264905, 18108370, 264629, 264908, 264909, 18108374, 18108377, 265008, 264910, 264637, 60170394, 264559, 265017, 264584, 264585, 264587, 264684, 264389
2880	94851439 (5759, 5760)	Novel Protein sim. GBank gl 4880703 gb AAD27741.1 AF13296 - (AF132966) CGI-32 protein [Homo sapiens]			264488, 52846365, 52846842, 22278994, 35898286, 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 35898052, 29331828, 264107, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 52844045, 264510, 265008, 264511, 265007, 264512, 265009, 264910, 264594, 21908754, 52846317, 52844288, 67188559, 264600, 264604, 264605, 264780, 264784, 264288, 264786, 264788, 264687, 264789, 21908768, 21908769, 35895917, 265021, 264690, 264692, 33657023, 33657349, 264629, 18108374, 35895855, 52845129, 33657109, 33657182, 27488282, 264634, 264635, 264636, 264637, 264638, 264557, 52844332, 264558, 264559, 83373044, 264404, 22278000, 264563, 264483, 264567, 264486
2881	87650539 (5761, 5762)	Novel Protein sim. GBank gl 733571 (U23452) - No definition line found [Caenorhabditis elegans]		UNCLASSIFIED	22278998, 29331822, 52644045, 21906765, 264639, 60432113
2882	87714367 (5763, 5764)	Novel Protein sim. GBank gl 1118112 (U41559) - No definition line found [Caenorhabditis elegans]			264488, 22278996, 22278999, 29331822, 29331826, 264908, 60170831, 60433358, 55812038, 264681, 264682, 264686, 264687, 264688, 21908768, 21908769, 264693, 263987, 18108374, 55811578, 58182323, 22279002, 264568
2883	95362875 (5765, 5766)	Novel Protein sim. GBank gl 488008 gb AAD31087.1 AF10693 - (AF106934) vitamin D receptor-interacting protein [Homo sapiens]	Contains protein domain (PF00400) - WD domain, G-beta repeat	UNCLASSIFIED	264112, 264682
2884	87784643 (5787, 5768)	Novel Protein sim. GBank gl 2224697 dbj BAA20832 - (AB002376) KIAA0378 [Homo sapiens]		UNCLASSIFIED	265018, 264634
2885	83006308 (5769, 5770)	Novel Protein sim. GBank gl 1255889 (U53344) - T07H6.5 gene product [Caenorhabditis elegans]	Contains protein domain (PF00084) - Sushi domain (SCR repeat)	complement receptor	60432049, 264259, 29331828, 284908, 264511, 264595, 60433438, 264598, 265017, 264605, 263989, 263972, 264555, 83373044, 87188518, 264566
2886	91237823 (5771, 5772)	Novel Protein sim. GBank gl 3882323 dbj BAA34521.1 - (AB018344) KIAA0801 protein [Homo sapiens]		UNCLASSIFIED	18108351, 264688, 264629, 264631, 264639, 83373044, 264482
2887	91227860 (5773, 5774)	Novel Protein sim. GBank gl 4508015 ref NP_003447.1 pZNF2 - zinc finger protein 205	Contains protein domain (PF00096) - Zinc finger, C2H2 type	transcript factor	264488, 264259, 29331828, 264508, 284906, 264593, 284758, 264766, 264769, 18108374, 83373044, 264486

2889	87606562 (5777, 5778)				UNCLASSIFIED	56994075, 28331824, 265009, 264760, 18108354, 264288
2890	78703853 (5778, 5780)	Novel Protein sim. GBank gl 854063 emb CAA58337 - (X83413) U88 [Human herpesvirus 6]			UNCLASSIFIED	264591, 264786
2891	80094428 (5781, 5782)	Novel Protein sim. GBank gl 3877750 emb CAB01508 - (Z78084) predicted using GeneFinder; similar to collagen; cDNA EST EMBL:D65885 comes from this gene; cDNA EST EMBL:D69451 comes from this gene; cDNA EST EMBL:D68028 comes from this gene; cDNA EST EMBL:D69658 comes from this gene...			UNCLASSIFIED	264591, 264595, 264369, 264685, 264693, 264628, 264583, 264568
2892	95419745 (5783, 5784)	Novel Protein sim. GBank gl 4929759 gb AAD34140.1 AF15180 - (AF151803) CGI-145 protein [Homo sapiens]			UNCLASSIFIED	65274572, 35696286, 29331828, 264110, 265008, 60433438, 265018, 265019, 18108351, 264288, 21906765, 21906766, 21906789, 55811957, 264690, 65274620, 263987, 35695855
2893	87788014 (5785, 5786)				UNCLASSIFIED	264259, 265010, 264682, 18108370, 264555, 264556
2894	87755985 (5787, 5788)	Novel Protein sim. GBank gl 5669015 gb AAD46135.1 - (AF080171) zinc finger protein ZNF232 [Homo sapiens]	Contains protein domain (PF00086) - Zinc finger, C2H2 type	transcript factor		264259, 265008, 60433438, 52644296, 265011, 264389, 35695917, 18108381, 18108382, 18108388
2895	86938778 (5789, 5790)	Novel Protein sim. GBank gl 3924708 emb CAA84646 - (Z35597) Weak similarity with sea squirt nidogen precursor protein (blastp score 71); cDNA EST EMBL:T02069 comes from this gene; cDNA EST EMBL:D78135 comes from this gene; cDNA EST EMBL:D73147 comes from this gene; cDNA EST EMB...	Contains protein domain (PF01437) - Plexin repeat			29331824, 265007, 264762, 264636, 264563
2896	87752122 (5791, 5792)	Novel Protein sim. GBank gl 4885549 ref NP_005456.1 ppK8G - protein kinase B gamma	Contains protein domain (PF00089) - Eukaryotic protein kinase domain	kinase		18108392, 18108394, 18108398, 22278998, 264259, 29331822, 28331824, 29148499, 264906, 264908, 265007, 265009, 265018, 265019, 264389, 264685, 264689, 21908766, 265021, 264693, 33657182, 264639, 18108384, 18108388, 264567
2897	85413057 (5793, 5794)	Novel Protein sim. GBank gl 4502877 ref NP_001296.1 pCLDN - Clostridium perfringens enterotoxin receptor 1			UNCLASSIFIED	60424179, 56182575, 22278996, 35696286, 22278997, 22278999, 60432049, 264259, 29331824, 29331825, 60424269, 60432289, 29331826, 29331828, 35696052, 264908, 56182435, 265009, 264910, 60170831, 60431735, 60433356, 60433438, 65274444, 55811386, 265018, 18108351, 264448, 264288, 264687, 21906765, 21906768, 21906767, 21906769, 55811957, 35695917, 264534, 33637023, 33857109, 35695783, 264628, 264629, 60431528, 18108374, 55810764, 55811576, 35696423, 35695855, 264555, 56182323, 18108385, 264404, 22279000, 22279002, 264568
2898	87750340 (5795, 5796)				UNCLASSIFIED	22278995, 264604, 18108385, 264566

2899	80357670 (5787, 5789)				UNCLASSIFIED	264764, 21806764, 264692
2900	84233538 (5789, 5800)	Novel Protein sim. GBank gl 4561470 emb CAB40137.1 - (Y18483) SLC7A9 protein [Homo sapiens]			glycoprotein	65274572, 56162575, 35696286, 60432049, 264259, 29331824, 66714117, 29331828, 35696052, 29331828, 66712502, 56182435, 35696052, 265007, 265009, 60433356, 264758, 265018, 264764, 264785, 264288, 264768, 21906764, 21906768, 21908769, 265020, 264692, 264693, 32833986, 264631, 83373044, 60432113
2901	87444731 (5801, 5802)	Novel Protein sim. GBank gl 4759272 ref NP_004614.1 pTTC4 - tetrairicopeptide repeat domain 4			phosphatase	22278995, 22278997, 22278999, 60432049, 29331822, 29331824, 29331825, 29331827, 35696052, 33658970, 264910, 265009, 21906754, 33657084, 87168474, 265010, 265018, 21906764, 21906765, 21906766, 21906767, 21906769, 33657023, 264693, 33657109, 33857349, 35696423, 35695855, 263981, 56182323, 22279002
2902	85745271 (5803, 5804)	Novel Protein sim. GBank gl 2414615 emb CAB16364 - (Z99259) hypothetical protein [Schizosaccharomyces pombe]				264683, 264691
2903	87606733 (5805, 5806)	Novel Protein sim. GBank gl 1079318 pir S52241 - XLCL2 protein - African clawed frog				264887, 22278994, 264259, 29331828, 29331828, 264905, 52644045, 56182435, 264511, 285017, 265018, 18108351, 264448, 264683, 264769, 264689, 35695917, 52644150, 87168518, 60432113, 22279002
2904	86458072 (5807, 5808)	Novel Protein sim. GBank gl 5639823 gb AAD45885.1 AF14367 - (AF143676) multispinning nuclear envelope membrane protein nrim [Homo sapiens]			UNCLASSIFIED	52646385, 22278999, 264259, 35696052, 265011, 265017, 284683, 264769, 35695917, 285020, 263987, 18108374, 35695855, 264637, 264952, 18108385, 18108387
2905	84449928 (5809, 5810)	Novel Protein sim. GBank gl 728837 sp P39194 ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII			oncogene	265009, 264681, 264682
2906	95341051 (5811, 5812)	Novel Protein sim. GBank gl 4689256 gb AAD27831.1 AF12185 - (AF121858) sorting nexin 8 [Homo sapiens]		Contains protein domain (PF00787) - PX domain	UNCLASSIFIED	22278998, 35696286, 22278998, 264259, 60432289, 29331828, 29331830, 66712502, 265009, 60170831, 33108954, 264448, 264683, 264288, 264689, 21906768, 21906767, 21906768, 55611957, 35695817, 265022, 52644150, 264691, 33657023, 264692, 264693, 35695855, 60432113, 264566

2807	91211383 (5813, 5814)	Novel Protein sim. GBank gji1707079 (U80451) - contains strong similarity to a DNA-J-like domain (PS:PS00636) [Caenorhabditis elegans]	Contains protein domain (PF00226) - eph DnaJ domain	52644507, 58182575, 58181686, 22278985, 56994075, 35698286, 80432049, 58182181, 35686052, 60431735, 264595, 55812038, 21908754, 55811386, 265019, 284682, 264369, 58181562, 21908766, 55811957, 35695917, 265020, 265021, 33657023, 33657109, 60431528, 55811578, 35686423, 35695855, 284638, 22279000
2808	80414246 (5815, 5816)	Novel Protein sim. GBank gji2873917 (AC002561) - putative ATP-dependent RNA helicase [Arabidopsis thaliana]	helicase	265009, 33109954, 18108351, 264766, 265021, 264691, 264692, 18108374, 264556, 284638, 264557, 264558
2809	67420225 (5817, 5818)		eph	264259, 87188474, 265018, 18108365, 264628
2810	86601075 (5819, 5820)	Novel Protein sim. GBank gji4539333[emb]CAB37483.1] - (AL035539) putative protein [Arabidopsis thaliana]		22278995, 264509, 264512, 265007, 33657402, 265017, 264369, 265022, 18108365, 284628
2811	94216615 (5821, 5822)	Novel Protein sim. GBank gji4469187[emb]CAB38415.1] - (AL031588) dJ1163.11.3 (novel protein similar to mouse B99) [Homo sapiens]	glucoamylase	52646365, 18108387, 22278995, 22278997, 22278998, 22278999, 28331824, 28331825, 52644045, 265008, 265018, 264448, 21908765, 21908766, 21908767, 21908768, 21908769, 55811957, 265021, 18108370, 18108372, 18108374, 22279000
2812	87731803 (5823, 5824)	Novel Protein sim. GBank gji4929837[gb]AAD34079.1[AF15184 - (AF151842) CGI-84 protein [Homo sapiens]	Contains protein domain (PF00904) - Involucrin repeat	52645156, 264082, 60432049, 284259, 52645080, 28331824, 28331825, 66712502, 33109954, 264760, 264683, 264288, 264686, 265021, 264693, 18108368, 283376, 284404
2813	87713823 (5825, 5826)	Novel Protein sim. GBank gji854065[emb]CAA58337] - (X63413) U88 [Human herpesvirus 8]	UNCLASSIFIED Zinc finger, C2H2 type	52644507, 52645156, 52646842, 58182575, 35698286, 22278997, 264259, 52645080, 28331827, 35696052, 28331828, 264628, 52644045, 58182435, 55812038, 52646317, 21908754, 52644286, 87168474, 265017, 265018, 265019, 18108351, 264682, 264686, 264689, 21908765, 21908766, 21908767, 21908768, 35695917, 285020, 52844150, 27486261, 27486262, 27486265, 35695763, 55811576, 35695855, 52644332, 22279000, 22279002, 264563
2814	87797300 (5827, 5828)			264537

2915	88081972 (5828, 5830)	Novel Protein sim. GBank gij5174485[re]NP_008030.1pK1AA - endocytic receptor (macrophage mannose receptor family)	Contains protein domain (PF00059) - eph Lectin C-type domain	264569, 264488, 264687, 264768, 21808769, 52646842, 21906767, 21908708, 56182375, 29148629, 35695917, 22278996, 22278997, 22278998, 265021, 22278999, 52644150, 284691, 284259, 60432049, 264692, 52645129, 33657109, 33657182, 29331827, 27486261, 35698052, 29331828, 27486262, 27486264, 27486265, 33657349, 29146498, 29146499, 264906, 264907, 18108370, 264908, 18108372, 52644045, 18108374, 56182435, 35695855, 284112, 284510, 265008, 60432229, 264593, 60433358, 56182323, 18108382, 55812038, 18108385, 33109954, 21908754, 33657084, 87168518, 265010, 265011, 60432113, 265017, 265018, 22279000, 265019, 55811150, 264681, 18108351, 264763, 284448, 264683, 264566, 18108354, 264389, 264288, 264768
2916	95337700 (5831, 5832)	Novel Protein sim. GBank gij5104851[dbj]BAA80165.1] - (AP000061) 305aa long hypothetical dTDP-4- dehydrohamnose reductase [Aeropyrum pernix]	dehydrogenase	32645156, 65274572, 22278994, 22278995, 35696288, 22278996, 22278997, 22278998, 22278999, 284259, 29331822, 29331824, 29331825, 66714117, 60432289, 29331826, 29331827, 29331828, 33656970, 264509, 264906, 29331830, 52644045, 264909, 56182435, 60170831, 264592, 264593, 33657402, 60433358, 52646317, 21808754, 33109954, 33657084, 52644298, 85858542, 265011, 265017, 265018, 265019, 18108351, 264448, 264288, 52644229, 21906765, 21906766, 21906767, 21908768, 21908769, 55811957, 35695917, 265021, 265022, 52644150, 33657023, 33657109, 33657182, 27486261, 27486262, 27486264, 35695783, 18108376, 55811578, 35696423, 85274791, 35695855, 52644332, 264557, 264638, 56182323, 18108387, 87168518, 22278002, 264482
2917	87454546 (5833, 5834)	Novel Protein sim. GBank gij3169065[emb]CAA19260.1] - (AL023704) putative translocation elongation factor-Tu fa mily [Schizosaccharomyces pombe]	UNCLASSIFIED	60433438, 264602, 264682, 87168518, 60432113
2918	85690529 (5835, 5836)	Novel Protein sim. GBank gij539218[pir]S38038 - hypothetical protein YKL201c - yeast (Saccharomyces cerevisiae)		264638
2919	87641497 (5837, 5838)	Novel Protein sim. GBank gij2564955 (AF030001) - unknown [Mus musculus]		66714117, 66712502, 263981

2020	87768523 (5839, 5840)				35896286, 22278897, 284258, 52645080, 28331824, 29331826, 29331827, 284828, 284909, 58182435, 284511, 284758, 33109954, 21908754, 52644298, 285010, 285011, 284801, 265017, 285019, 284681, 284687, 21906767, 265021, 52644150, 284690, 284691, 284692, 284693, 33657109, 33657182, 27488282, 27488284, 27488285, 35898423, 35895855, 284832, 284636, 284637, 284638, 58182323, 60170394, 18108385, 87168518, 60432113
2021	91639982 (5841, 5842)	Novel Protein sim. GBank gi 4580013 gb AAD24202.1 U83194_ - (U83194) TRAF4-associated factor 2 (Homo sapiens)	Contains protein domain (PF00787) - PX domain		35896286, 22278897, 284258, 52645080, 28331824, 29331826, 29331827, 284828, 284909, 58182435, 284511, 284758, 33109954, 21908754, 52644298, 285010, 285011, 284801, 265017, 285019, 284681, 284687, 21906767, 265021, 52644150, 284690, 284691, 284692, 284693, 33657109, 33657182, 27488282, 27488284, 27488285, 35898423, 35895855, 284832, 284636, 284637, 284638, 58182323, 60170394, 18108385, 87168518, 60432113
2022	87740762 (5843, 5844)	Novel Protein sim. GBank gi 4589514 db BAA76779.1 - (AB023152) KIAA0935 protein (Homo sapiens)	Contains protein domain (PF01074) - Glycosyl hydrolases family 38	kinase	284094, 284259, 29331822, 29331824, 28331826, 29331827, 35896052, 29148498, 284104, 284105, 284107, 284509, 284110, 284112, 284512, 60433358, 21908754, 87168474, 265017, 18108351, 284288, 21908765, 21908766, 21908767, 21908769, 35895917, 265021, 263974, 18108374, 283978, 283977, 18108376, 284555, 283981, 58526488, 87168518, 22279000, 22279002, 284808, 284909, 284511, 265008, 285008, 284593, 33857402, 60174639, 18108351, 284763, 21908765, 29148627, 35895917, 284692, 284629, 263978, 5811578, 35895855, 284555, 284558, 58182323, 60170394, 22279000, 284486
2023	95337789 (5845, 5846)	Novel Protein sim. GBank gi 4835268 emb CAB42898.2 - (Z83844) dJ37E16.4 (similar to mouse p116Rip protein) (Homo sapiens)	Contains protein domain (PF00169) - PH domain	struct	284488, 18108397, 22278995, 22278996, 22278997, 22278998, 22278999, 29331825, 29331826, 29331827, 29331830, 284511, 285008, 33857402, 285011, 285017, 265018, 284883, 18108354, 21906785, 21908767, 21908768, 21908769, 52644150, 284691, 284692, 33857109, 263974, 18108376, 284631, 284636, 18108385, 18108387, 22279000, 284563, 284588
2024	87791987 (5847, 5848)	Novel Protein sim. GBank gi 2130395 pir J57234 - ribosomal protein L36, mitochondrial - yeast (Saccharomyces cerevisiae)	Contains protein domain (PF00444) - Ribosomal protein L36	ribosomalprot	22279000, 284563, 284588, 265017, 284628, 20281152, 284558
2025	95080120 (5849, 5850)	Novel Protein sim. GBank gi 2388986 emb CAB11718 - (Z98980) actin associated protein [Schizosaccharomyces pombe]	UNCLASSIFIED		58182575, 35898286, 284258, 60432289, 29331827, 284508, 52644045, 284910, 284591, 60432228, 55812038, 21908754, 284681, 284448, 264683, 264288, 284685, 52644229, 284689, 21906785, 21908768, 21908769, 285021, 285022, 60170615, 284692, 33657023, 284693, 33657109, 35898423, 65274781, 58182323

2928	85343003 (5851, 5852)	Novel Protein sim. GBank gl 283032 pir S22456 - hydroxyproline-rich glycoprotein - perennial leosinte			29331828, 265011, 264768, 264689, 264764, 264288, 264630, 264637
2927	80408018 (5853, 5854)				
2928	20432179 (5855, 5856)	Novel Protein sim. GBank gl 3413320 emb CAA06915 - (AJ008215) CMP-N-acetylneuraminic acid synthetase [Mus musculus]		UNCLASSIFIED	264559
2929	81622820 (5857, 5858)			UNCLASSIFIED	264569, 264488, 22278994, 35686286, 22278998, 22278999, 264259, 264094, 264259, 52645080, 29331822, 29331824, 66714117, 29331825, 29331826, 29331827, 35686052, 33856970, 264109, 29331830, 52644045, 265009, 33109954, 52644286, 87168559, 264760, 264762, 264448, 264764, 264288, 264766, 264768, 21908765, 21908768, 21908769, 35685917, 264891, 33657023, 264693, 33657109, 18108374, 263976, 35698423, 35695855, 263981, 22279000, 22278002, 264587, 264486
2930	95302755 (5859, 5860)			UNCLASSIFIED	56182575, 56181686, 35698286, 22278996, 22278998, 22278999, 264259, 29331825, 60432289, 29331828, 264905, 52644045, 56182435, 265009, 60170831, 264592, 60432228, 60433356, 87168474, 265010, 265011, 265017, 265018, 265019, 264762, 264448, 264683, 264288, 264768, 21908765, 21908769, 35695917, 60170815, 33857023, 33857109, 264628, 18108370, 18108372, 35698423, 35695855, 264556, 56182323, 60432113, 264587
2931	94312693 (5861, 5862)	Novel Protein sim. GBank gl 3768433 (AF098505) - similar to Arabidopsis thaliana male sterility protein 2 (SW:Q08891) [Caenorhabditis elegans]	Contains protein domain (PF00471) - Ribosomal protein L33	UNCLASSIFIED	52645156, 22278997, 22278998, 29331822, 52645080, 29331824, 60432289, 33856970, 60433356, 60433438, 33109954, 21908765, 21908768, 21908767, 21908768, 265020, 52644150, 33857023, 33857109, 33857182, 27486265, 35686423, 35695855, 264555, 87168518, 60432113, 264568
2932	79632623 (5863, 5864)				264906, 264907
2933	91720776 (5865, 5866)	Novel Protein sim. GBank gl 3378058 (AF017777) - helicase [Drosophila melanogaster]	helicase		264488, 18108392, 56182575, 22278999, 264091, 264259, 29331825, 60432289, 29331827, 264508, 52644045, 56182435, 265007, 265009, 264592, 60433356, 60433438, 21908754, 265017, 264682, 264288, 52644228, 21908765, 21908768, 21908769, 265022, 52644150, 33857023, 33657109, 27486265, 264635, 264636, 60170394, 56182323, 18108385, 60432113, 264585, 264568, 264567

2834	86576025 (5867, 5868)				22278997, 22278999, 29331824, 33657402, 264691, 27486282, 264628, 87168518, 22278000
2835	86410579 (5869, 5870)			UNCLASSIFIED	58182575, 22278995, 60433358, 33657402, 264758, 33109954, 21808754, 265018, 265019, 284448, 264769, 21808784, 21808765, 265021, 264692, 33657023, 33657109, 33657349, 55810784, 22278000
2836	87605863 (5871, 5872)	Novel Protein sim. GBank gi4153862 (AC005065) - determined by GENSCAN prediction and spliced EST; match to EST R84329 (NID:942735) [Homo sapiens]	Contains protein domain (PF00856) - SET domain	nuclease	22278997, 29331827, 29331828, 265009, 265017, 264605, 265020, 55811578, 18108387, 60432113, 284563
2837	94853096 (5873, 5874)	Novel Protein sim. GBank gi5174409[ref NP_006101.1 pCD28 - CD2 antigen (cytoplasmic tail)-binding protein 2		UNCLASSIFIED	56994075, 22278999, 284259, 60432049, 29331822, 56182181, 29331827, 29331828, 264908, 264908, 264909, 56182435, 265008, 284512, 264910, 60170831, 60433358, 265011, 265018, 18108351, 284448, 264288, 284766, 52844228, 21908765, 29148784, 65274791, 284556, 56182323, 60170394, 284558, 60432113, 284565, 264486, 284587
2838	95419773 (5875, 5876)	Novel Protein sim. GBank gi3319990[emb CAA76720] - (Y17267) ubiquitin-conjugating enzyme [Mus musculus]	Contains protein domain (PF00179) - Ubiquitin-conjugating enzyme	ubiquitin	22278997, 22278998, 22278999, 284480, 284259, 29331822, 29331824, 66714117, 29331827, 35696052, 264107, 284805, 66712502, 52844045, 56182435, 284511, 285008, 265009, 60432229, 33657402, 60433438, 55812038, 21808754, 8568842, 285010, 285011, 87168559, 265017, 265018, 265019, 264681, 264288, 264689, 21908765, 21908787, 21908788, 55811957, 35695917, 265020, 60170615, 264690, 264691, 264692, 33657023, 284693, 65274620, 33657109, 18108370, 18108374, 263978, 35696423, 35695855, 264555, 284556, 18108381, 56182323, 60170394, 83373044, 18108385, 58528486, 60432113, 22278002
2839	87786622 (5877, 5878)	Novel Protein sim. GBank gi3979900[emb CAA99909] - (Z75547) similar to WD domain, G-beta repeat; cDNA EST yk371b7.5 comes from this gene; cDNA EST yk312h1.5 comes from this gene; cDNA EST yk465d5.5 comes from this gene; cDNA EST yk472c4.5 comes from this gene; cDNA EST yk292f8....	Contains protein domain (PF00400) - WD domain, G-beta repeat	ATPase-associated	264807, 265018, 264681, 264685, 264686

2940	95011103 (5879, 5880)			UNCLASSIFIED	22278986, 29331822, 29331824, 66714117, 29331826, 29331828, 264905, 264908, 66712502, 29331830, 265011, 265017, 284784, 264369, 21906766, 21906767, 33657023, 33657109, 32833966, 18108374, 18108377, 264634, 83373044, 18108385, 18108387, 264566, 284557
2941	21423370 (5881, 5882)	Novel Protein sim. GBank gij3413872[dbj]BAA32300] - (AB007924) KIAA0455 protein [Homo sapiens]		UNCLASSIFIED	
2942	87430203 (5883, 5884)	Novel Protein sim. GBank gij1172845[sp]P46829[R825_RABIT - RAS-RELATED PROTEIN RAB-25]		glycoprotein	284910, 265010, 264768
2943	95314504 (5885, 5886)	Novel Protein sim. GBank gij4929653[gb]AAD34087.1[AF15185] - (AF151850) CGI-92 protein [Homo sapiens]		collagen	60432049, 264259, 60432288, 29331827, 29148498, 265008, 264593, 60433356, 60433438, 265010, 265011, 265017, 265018, 284683, 264766, 18108381, 65274727, 60432113, 264567
2944	95081063 (5887, 5888)	Novel Protein sim. GBank gij4678282[emb]CAB41180.1] - (AL049880) t-acylcort-3-phosphate acyltransferase-like protein [Arabidopsis thaliana]	Contains protein domain (PF00415) - ATPase associated Regulator of chromosome condensation (RCC1)		56994075, 22278988, 60432049, 264259, 29331822, 29331824, 60424268, 60432289, 29331828, 29331828, 264905, 264907, 52844045, 264909, 264511, 265006, 265009, 284594, 21906754, 87168559, 264603, 265017, 265018, 18108351, 264682, 264766, 264687, 264689, 21908765, 21908766, 21906767, 21908768, 21906769, 265021, 60170615, 52644150, 264690, 264691, 33657023, 264692, 264693, 33657109, 33657182, 33657349, 18108370, 18108374, 18108377, 55811578, 35698423, 35698855, 264635, 264555, 264556, 56182323, 60170394, 264558, 264559, 83373044, 56526486, 87168518, 60432113, 22279002, 264482, 264563, 264484, 284567
2945	94233560 (5889, 5890)	Novel Protein sim. GBank gij728631[sp]P39188[ALU1_HUMAN - IIII ALU SUBFAMILY J WARNING ENTRY IIII]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	UNCLASSIFIED	60424179, 22278995, 22278996, 22278998, 22278999, 264259, 56182181, 29331824, 60424269, 60432289, 35696052, 264808, 265006, 60433356, 55812038, 264759, 5581388, 265018, 264681, 18108351, 264448, 264683, 264369, 264288, 264687, 56181562, 21906767, 21908768, 21908769, 35695917, 265020, 265021, 264693, 60431528, 55810784, 35698423, 35698855, 264630, 60170394, 83373044, 22279000, 264566, 284567

2946	84317315 (5891, 5892)	Novel Protein sim. GBank gi 5441952 gb AAD43185.1 AF07288 - (AF072884) peroxisomal membrane protein PMP 24 [Homo sapiens]		UNCLASSIFIED	264488, 264259, 264508, 264509, 264908, 264907, 264809, 264510, 264511, 265007, 264512, 264810, 264591, 264593, 18108351, 264764, 264288, 264884, 264769, 265021, 264892, 33657109, 264828, 264828, 18108374, 264631, 264634, 264636, 264637, 18108380, 264638, 264639, 83373044, 264555, 264586, 264488, 264587
2947	87362852 (5893, 5894)	Novel Protein sim. GBank gi 3540281 gb AAC34383.1 - (AF058116) All-1 related protein [Fugu rubripes]		UNCLASSIFIED	22278995, 22278998, 22278997, 22278999, 29146498, 264508, 29331830, 265007, 265008, 265009, 60432228, 21908754, 265010, 265017, 265019, 264766, 264685, 21908765, 21908766, 21908767, 21908768, 21908769, 265020, 264628, 18108370, 264829, 264630, 18108387, 60432113
2948	87626527 (5895, 5896)	Novel Protein sim. GBank gi 5568614 gb AA865654.2 - (AF001533) mitogen-induced [Mus musculus]			52646842, 22278995, 264259, 29331824, 28331825, 29331827, 29331830, 264809, 265007, 265009, 265019, 264763, 264684, 264288, 264685, 264686, 21908767, 264691, 264692, 264693, 18108374, 55811576, 18108385, 22279002, 264593, 264587
2949	88175545 (5897, 5898)	Novel Protein sim. GBank gi 2132923 pir J567133 - probable membrane protein YOR240w - yeast (Saccharomyces cerevisiae)		UNCLASSIFIED	22278998, 22278997, 60432289, 29331826, 29331827, 29331828, 35698052, 29146499, 264104, 264107, 264905, 68712502, 264908, 60433358, 60433438, 87188558, 264764, 52644229, 58181582, 21908767, 21908768, 21908769, 265022, 60170815, 33657023, 35696423, 263981, 264558, 60432113, 22279002
2950	95086870 (5899, 5900)	Novel Protein sim. GBank gi 468102 sp P34629 YQJ8. CAEEL - PUTATIVE AMINOPEPTIDASE ZK353.6 IN CHROMOSOME III	Contains protein domain (PF00883) - Cytosol aminopeptidase family	peptidase	264488, 35698286, 264259, 35696052, 264907, 265007, 264910, 265017, 265018, 264288, 264768, 35695917, 265020, 18108382, 18108370, 18108378, 35696423, 65274791, 35695855, 264558, 56528486, 264486
2951	87392357 (5901, 5902)	Novel Protein sim. GBank gi 4688902 emb CAB41450.1 - (AJ238248) centaurin beta2 [Homo sapiens]			264693

2952	95329952 (5903, 5904)	Novel Protein sim. GBank gij5596693[emb](CAB51405.1) - (AL098881) hypothetical protein [Homo sapiens]	Contains protein domain (PF00650) - CRAL/TRIO domain.		264887, 52645156, 21906766, 21906769, 22278996, 265020, 264890, 60432049, 264259, 264693, 29331822, 18108365, 29331825, 60432289, 33657109, 18108368, 29331827, 35896052, 27486262, 264508, 264905, 20281149, 264908, 264807, 29331830, 264908, 264909, 35895855, 264511, 265008, 265009, 264910, 264635, 264636, 60432229, 264638, 60433356, 264639, 264756, 87188518, 265017, 22279000, 22279002, 264760, 264563, 264482, 18108351, 264448, 264288
2953	86093575 (5905, 5906)	Novel Protein sim. GBank gij149522[sp]P10658[SERC_RABIT - PROBABLE PHOSPHOSERINE AMINOTRANSFERASE (PSAT) (ENDOMETRIAL PROGESTERONE-INDUCED PROTEIN) (EPIP)]	Contains protein domain (PF00266) - Aminotransferases class-V	UNCLASSIFIED	18108398, 56984075, 22278996, 29331822, 29331824, 29331825, 29331827, 35896052, 264508, 264905, 264908, 264907, 264510, 264591, 264594, 33857402, 264595, 264596, 264758, 52646317, 21906754, 33857084, 52644296, 87188559, 264600, 264760, 264691, 18108351, 264764, 264399, 264288, 264687, 21908765, 21908766, 21908767, 21908768, 21908769, 35895917, 33857023, 18108364, 52645129, 33857109, 33857349, 18108374, 263978, 35895855, 264637, 264638, 87188518, 264482, 264563, 264565
2954	86086288 (5907, 5908)	Novel Protein sim. GBank gij4865261[ref]NP_005251.1[pGDF9 - growth differentiation factor 9]	Contains protein domain (PF00019) - Transforming growth factor beta like domain	Ig1	29331822
2955	87698426 (5909, 5910)	Novel Protein sim. GBank gij3452473 (AF084205) - serine/threonine protein kinase TAO1 [Rattus norvegicus]	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	kinase	264259, 29331822, 29331824, 29331825, 29331826, 35896052, 264808, 52644045, 264512, 60432229, 265018, 265019, 55811150, 264789, 21908767, 21908768, 21906789, 265021, 60170615, 55810764, 264567
2956	85789745 (5911, 5912)	Novel Protein sim. GBank gij4889254[gb]AAD27830.1[AF121857] sorting nexin 7 [Homo sapiens]	Contains protein domain (PF00787) - PX domain		22278996, 264259, 29331827, 264808, 21908788
2957	90933301 (5913, 5914)	Novel Protein sim. GBank gij4503023[ref]NP_000089.1[pCPT2 - carnitine palmitoyltransferase II precursor]		cadherin	22278999, 264259, 29331824, 29331827, 265008, 264595, 264758, 265010, 265011, 264448, 264763, 264683, 264288, 264685, 18108357, 29148629, 264690, 18108362, 264693, 18108370, 60431528, 18108374, 264634, 18108381, 56182323, 18108382, 18108385, 18108388, 56526486, 87188518, 264482, 264487
2958	87440014 (5915, 5916)	Novel Protein sim. GBank gij4240257[db]BAA74907.1 - (AB020691) KIAA0884 protein [Homo sapiens]			264595, 264596, 264681, 264369, 264629, 264631, 264587

2859	95100420 (5917, 5918)	Novel Protein sim. GBank gij988221 (U33005) - Tbc1 [Mus musculus]	Contains protein domain (PF00566) - oncogene TBC domain	283994, 22278997, 264259, 60432049, 29331826, 29331828, 35696052, 29331830, 68712502, 56182435, 265008, 264512, 265008, 265009, 60433358, 60433438, 264596, 265017, 265018, 264683, 264288, 264766, 264769, 21906766, 21906767, 21906769, 265020, 60170815, 264692, 27466285, 18108374, 85274791, 35695855, 83373044, 56526486, 60432113
2960	87420091 (5918, 5920)		UNCLASSIFIED	35696286, 56182435, 87168474, 265010, 60170815, 35698423, 56182323, 18108383, 87168518, 264483
2961	85413416 (5921, 5922)	Novel Protein sim. GBank gij5596646[emb]CAB05177.2] - (282266) predicted using GeneFinder; similar to WD domain, G-beta repeats [Caenorhabditis elegans]	Contains protein domain (PF00400) - WD domain, G-beta repeat	22278997, 22278999, 264259, 29331822, 29331824, 29331826, 29331828, 264907, 264908, 52644045, 265008, 33657402, 21906754, 87168474, 265011, 87168559, 265017, 21906769, 265020, 60170815, 264692, 33657023, 35695763, 18108370, 18108374, 35698423, 264632, 264636, 18108385, 87168518, 22279002, 284584, 264567
2962	87912700 (5923, 5924)		UNCLASSIFIED	35696286, 22278997, 264092, 264094, 264259, 29331824, 66714117, 29331825, 60432289, 29331826, 29331827, 29331828, 35698052, 264508, 264905, 264509, 264907, 264908, 264909, 264510, 264512, 264593, 264594, 60433438, 264758, 52646317, 264602, 264603, 264605, 264760, 264762, 264764, 264288, 264768, 264688, 264768, 264769, 35695917, 265020, 264691, 264634, 264636, 264637, 264638, 264639, 18100385, 264593, 264565, 264586, 264587, 264486
2963	95313464 (5925, 5926)	Novel Protein sim. GBank gij4240223[dbj]BAA74890.1] - (AB020674) KIAA0887 protein [Homo sapiens]	Contains protein domain (PF00010) - Helix-loop-helix DNA-binding domain	18108392, 56994075, 22278998, 22278999, 29331822, 29331825, 29331826, 29331827, 29331828, 265007, 265008, 264592, 264594, 21906754, 265018, 264760, 264687, 2646827, 29148784, 265020, 33657023, 264693, 65274620, 33657182, 27466281, 264629, 55810764, 35696423, 264555, 264636, 264637, 264557, 264558, 264563
2964	94324617 (5927, 5928)		UNCLASSIFIED	264259, 29331828, 33657402, 265017, 265018, 264692, 18108386, 35698423, 83373044, 18108388

2965	80384762 (5929, 5930)	Novel Protein sim. GBank gij4885447jrefjNP_005432.1 pkRML - Kreister (mouse) mal- related leucine zipper homolog			Transcript factor	264259, 29331828, 264508, 264509, 264905, 264907, 264908, 264909, 264511, 265008, 264910, 264591, 264593, 264594, 33657402, 265011, 264760, 264762, 264764, 264288, 264685, 264766, 264692, 33657109, 264628, 264629, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264637, 264638, 264639, 264563, 264567, 18108391 60432289, 264682, 264448
2966	81725248 (5931, 5932)	Novel Protein sim. GBank gij528275 emb CAB45690.1 - (AJ243177) Xenopus RPA Interacting protein alpha [Xenopus laevis]				
2967	94658303 (5933, 5934)	Novel Protein sim. GBank gij624225 (U19181) - Rabin3 [Rattus norvegicus]			UNCLASSIFIED	264488, 264508, 264509, 264908, 264909, 264511, 264910, 264594, 264758, 85858542, 264762, 264764, 265021, 264556, 18108381, 264584, 264486
2968	95302776 (5935, 5936)	Novel Protein sim. GBank gij4929715 gb AAD34118.1 AF15188 - (AF151881) CGI-123 protein [Homo sapiens]	Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger)			264687, 52845156, 21908765, 52846365, 21908767, 18108398, 35696423, 22278996, 35698286, 22278997, 265020, 22278999, 265021, 265022, 264093, 264636, 264680, 52844150, 264259, 33657023, 52845080, 264693, 29331822, 56182181, 29331824, 68714117, 29331825, 33109954, 52845129, 29331826, 21908754, 33657182, 29331827, 29331828, 35698052, 27486282, 87168518, 87168474, 265010, 87168559, 265018, 22279000, 265019, 22279002, 264563, 18108351, 264908, 264907, 264448, 68712502, 264568, 264369, 264288 52846842, 22278996, 22278998, 22278999, 60432049, 264259, 29331824, 29331825, 29331826, 29331828, 264509, 264909, 52844045, 56182435, 265009, 60433438, 55812038, 21908754, 265011, 87168559, 265018, 265019, 264448, 264288, 264369, 52844229, 21908768, 21908769, 21908769, 29148784, 265020, 265021, 52844150, 264691, 33657109, 18108374, 56182323, 60170394, 87168518, 60432113, 22279000
2969	95310957 (5937, 5938)	Novel Protein sim. GBank gij3024743 sp O24734 THSA_SULS7 - THERMOSOME, ALPHA SUBUNIT (CHAPERONIN ALPHA SUBUNIT)		eph		
2970	88088071 (5939, 5940)	Novel Protein sim. GBank gij3165407 (AC004755) - fos37502_1 [Homo sapiens]	Contains protein domain (PF00046) - Homeobox domain	homeobox		

2871	94196830 (5941, 5942)	Novel Protein sim. GBank glf728837isp/p39194/ALU7 SQ WARNING ENTRY IIII		lm7	264488, 56182575, 35686286, 56894075, 29331824, 29331826, 28146499, 284508, 284905, 264907, 284112, 284910, 21908754, 87168559, 285018, 285019, 18108351, 284889, 21908765, 21908767, 21908768, 265020, 265021, 60170815, 18108364, 284628, 284629, 18108374, 284638, 284556, 284558, 83373044, 18108384, 18108385, 87168518, 284584, 284587
2872	86825943 (5943, 5944)	Novel Protein sim. GBank glf728836isp/p39193/ALU6_HUMAN - IIII ALU SUBFAMILY SP WARNING ENTRY IIII		kinase	265017, 35695917, 265021, 33657109, 22278002, 264583
2873	91215301 (5945, 5946)	Novel Protein sim. GBank glf2746789 (AF040842) - No definition line found [Caenorhabditis elegans]		UNCLASSIFIED	29331822, 264692, 33657349, 55811576, 264583
2874	91873002 (5947, 5948)	Novel Protein sim. GBank glf768117 (L41834) - nuclear protein [Ensis minor]		UNCLASSIFIED	18108392, 52644507, 56182575, 58181686, 22278995, 22278996, 35686286, 22278997, 22278998, 22278999, 264259, 52645080, 29331824, 29331825, 66714117, 60424269, 29331826, 29331827, 29331828, 35696052, 68712502, 284808, 52644045, 265007, 284910, 285009, 60433438, 33108954, 21908754, 55811386, 52644288, 87168474, 87168559, 265017, 265018, 265019, 18108351, 264448, 264389, 264288, 52644229, 18108359, 21908765, 21908767, 21908768, 35695917, 265020, 265021, 52644150, 264691, 264692, 33657023, 27488282, 27486284, 35695763, 18108370, 18108376, 55810764, 55811576, 35698423, 35695855, 264630, 264635, 264557, 52644332, 264558, 83373044, 18108387, 87168518, 60432113, 22279000, 264482, 264487
2875	95326213 (5949, 5950)	Novel Protein sim. GBank glf3880812[emb]CAA19508] - (AL023839) similar to HECT-domain (ubiquitin-transferase).; cDNA EST yk480d10.5 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00632) - HECT-domain (ubiquitin- transferase).	ubiquitin	29331824, 29331827, 29331828, 264910, 83658542, 265011, 265018, 264448, 264288, 264769, 21908767, 265020, 264691, 264559, 83373044
2876	87771202 (5951, 5952)	Novel Protein sim. GBank glf5679136[gb]AAD46874.1[AF160934] - (AF160934) BcDNA.LD14189 [Drosophila melanogaster]		transport	22278998, 264808, 265007, 265010, 265011, 265017, 265019, 18108351, 284685, 284689, 18108370, 264639, 18108385
2877	91725254 (5953, 5954)	Novel Protein sim. GBank glf5282751[emb]CAB45680.1] - (AJ243177) Xenopus RPA interacting protein alpha [Xenopus laevis]		UNCLASSIFIED	284509, 264288

2978	87332059 (5955, 5958)	Novel Protein sim. GBank gij746549 (U23522) - No definition line found [Caenorhabditis elegans]	Contains protein domain (PF00480) - ROK family	UNCLASSIFIED	22278995, 22278996, 22278997, 22278999, 284259, 60432289, 29331827, 29146499, 56182435, 265006, 265007, 265009, 60433358, 60433438, 21908754, 265010, 265011, 265017, 265018, 265019, 264288, 284685, 284688, 21908765, 21908768, 21908767, 21908768, 21908769, 265020, 265021, 265022, 35696423, 264639, 60432113, 22279000, 22279002
2979	91725258 (5957, 5958)	Novel Protein sim. GBank gij5282751 (emb CAB45690.1 - (AJ243177) Xenopus RPA Interacting protein alpha [Xenopus laevis]	complement		264488, 65274572, 56994075, 22278999, 264093, 29331822, 29331824, 264288, 55811957, 33657023, 33657109, 18108370, 55811576, 56182323, 60432113, 264482, 265009, 21908767, 263981, 22279000
2980	86286600 (5959, 5960)			UNCLASSIFIED	264629, 264564
2981	87376330 (5961, 5962)				
2982	95303675 (5963, 5964)	Novel Protein sim. GBank gij4929787 (gb AAD34144.1 AF15190 - (AF151907) CGI-149 protein [Homo sapiens]			22278995, 56994075, 22278996, 22278997, 22278998, 22278999, 264092, 26331824, 29331827, 29331828, 264905, 264591, 284592, 264594, 264595, 264596, 33657084, 264448, 21908765, 21908768, 21908767, 21908768, 21908769, 265020, 265022, 18108365, 33657182, 33657349, 35696423, 83373044, 22279000, 22279002
2983	91725258 (5965, 5968)	Novel Protein sim. GBank gij5262751 (emb CAB45690.1 - (AJ243177) Xenopus RPA Interacting protein alpha [Xenopus laevis]			60424178, 52646842, 18108398, 22278997, 264093, 60432049, 264259, 29331822, 60432289, 33656970, 264905, 52644045, 265006, 60431735, 87168474, 265018, 265019, 18108351, 264448, 21908765, 21908768, 35695917, 33657023, 52845129, 18108370, 35696423, 83373044, 56526486, 60432113, 264404, 22279002
2984	94136487 (5967, 5968)	Novel Protein sim. GBank gij293734 (AC002542) - similar to C. elegans F11A10.5; 80% similarity to Z68297 (PID:g1130619) [Homo sapiens]	ATPase-associated		
2985	87095072 (5969, 5970)	Novel Protein sim. GBank gij103160 (pir S22126 - finger protein unkempt - fruit fly (Drosophila melanogaster)	UNCLASSIFIED		264910, 55812038, 56181562, 55811957, 264628, 55810764, 264632, 264635, 60432113
2986	86284861 (5971, 5972)				55811957, 264566
2987	86455934 (5973, 5974)		UNCLASSIFIED		264369

2988	95357753 (5975, 5976)	Novel Protein sim. GBank gij4678028igbAAAD27002.11 - (AF077207) HSPC021 (Homo sapiens)		UNCLASSIFIED	264488, 65274572, 22278995, 22278998, 22278997, 22278999, 264092, 264094, 264259, 60432049, 29331824, 29331826, 60432288, 35696052, 29331828, 264107, 264905, 264807, 264908, 66712502, 264828, 264909, 58182435, 265006, 265007, 265008, 60170831, 60432229, 264593, 60433358, 264757, 60433438, 21908754, 265010, 265011, 87168559, 265017, 265018, 264682, 264448, 264369, 264288, 264685, 52844229, 21908765, 21906787, 21908789, 35695917, 265021, 265022, 52644150, 264690, 33657023, 65274620, 263967, 33657109, 27486262, 18108370, 18108372, 18108374, 55810764, 65274791, 35695855, 264635, 264636, 264637, 263981, 264638, 56182323, 83373044, 60432113, 22279000, 264563, 264584, 264565, 264566, 264567
2989	91225118 (5977, 5978)	Novel Protein sim. GBank gij113871spIP23984IALUF_HUMAN - IIII ALU CLASS F WARNING ENTRY IIII	kinase		22278998, 22278997, 264905, 264511, 60170831, 264593, 265019, 21908765, 21906787, 21908768, 18108374
2990	87330444 (5979, 5980)	Novel Protein sim. GBank gij2829836spIP97348RHOD_MOUSE - RHO-RELATED GTP-BINDING PROTEIN RHOD	Contains protein domain (PF00071) - Ras family	oncogene	265007, 264512, 18108351, 264288, 264689, 265020, 264691, 33657023, 33657109
2991	84325361 (5981, 5982)			UNCLASSIFIED	264563
2992	85425164 (5983, 5984)			UNCLASSIFIED	264259, 265019, 264689, 18108385
2993	94325363 (5985, 5986)			UNCLASSIFIED	264488, 29331822, 265017, 264761, 21906769, 65274791, 263981, 264565
2994	94136634 (5987, 5988)	Novel Protein sim. GBank gij2488548spIO50659YU02_MYCTU - HYPOTHETICAL 28.7 KD PROTEIN CY339.02	transport		22278994, 22278995, 56994075, 22278997, 22278999, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 265008, 265009, 264910, 33109954, 87168474, 87168559, 265018, 265019, 264448, 264286, 21906766, 21906767, 21908768, 21906769, 265021, 265022, 33657023, 264693, 35695855, 83373044, 18108385, 22279000, 264565, 264586, 264905, 264907, 265019, 18108351, 264683
2995	875-1070 (5989, 5990)	Novel Protein sim. GBank gij2734081 (AF000195) - similar to oxysterol-binding proteins [Caenorhabditis elegans]			85274572, 35696286, 264259, 29331824, 35696052, 29148499, 264508, 264907, 265007, 265008, 60433438, 18108348, 265017, 264681, 264683, 264288, 264768, 264769, 264689, 35695917, 60170815, 33657023, 264692, 264634, 264555, 18108381, 18108382, 18108386, 264484
2996	91013788 (5991, 5992)	Novel Protein sim. GBank gij2628912 (AC002291) - Similar ATP-dependent RNA Helicase [Arabidopsis thaliana]	Contains protein domain (PF00270) - DEAD/DEAH box helicase	helicase	

2997	87627440 (5993, 5994)	Novel Protein sim. GBank gi 4589652 dbj BAA78848.1 - (AB023221) KIAA1004 protein [Homo sapiens]		homeobox	264488, 56182575, 264259, 66714117, 29331828, 35898052, 264508, 264509, 264907, 264908, 265006, 87168474, 265019, 264448, 264682, 264685, 264766, 21908764, 21908766, 21908768, 21908769, 27486281, 18108374, 35898423, 264634, 264635, 264638, 264557, 18108385, 87168518
2998	86095381 (5995, 5996)	Novel Protein sim. GBank gi 3947599 emb CAA22252 - (AL034384) cDNA EST yk255b9.3 comes from this gene; cDNA EST yk255b9.5 comes from this gene; cDNA EST EMBL:M75923 comes from this gene [Caenorhabditis elegans]		UNCLASSIFIED	52646365, 22278997, 264508, 264908, 18108351, 21908765, 21908767, 18108370, 18108374, 35898423, 264638, 264639
2998	94847055 (5997, 5998)	Novel Protein sim. GBank gi 15408 sp P18835 CC19_CAEEL - CUTICLE COLLAGEN 19		UNCLASSIFIED	56182575, 22278996, 28147620, 29331825, 29146498, 28146499, 264905, 66712502, 265006, 265009, 21906754, 85858542, 18108351, 29148627, 29148629, 60170615, 33657109, 27486262, 18108370, 18108374, 264558, 264557, 264558, 60170394, 18108385, 264563
3000	95099370 (5999, 6000)	Novel Protein sim. GBank gi 11631174 (U32575) - similar to yeast Sec6p, Swiss-Prot Accession Number P32844; similar to mammalian B94, Swiss-Prot Accession Number Q03169; Method: conceptual translation supplied by author [Rattus norvegicus]		UNCLASSIFIED	264887, 22278997, 22278999, 264259, 29331822, 29331824, 35898052, 29146498, 264508, 264905, 264908, 264907, 264908, 264908, 264510, 264511, 265006, 265007, 265008, 265009, 264910, 33657402, 284757, 264595, 264598, 264758, 21906754, 265011, 264600, 265017, 265018, 264605, 265019, 264760, 264761, 264762, 264681, 264682, 264764, 264288, 264685, 264766, 264686, 264768, 264769, 21906765, 21906768, 35895917, 265020, 264691, 264692, 33657023, 264693, 33657109, 33657182, 27486261, 264628, 264629, 18108374, 18108376, 35898423, 35895855, 264630, 264631, 264632, 264634, 264635, 264636, 264637, 264556, 264638, 264639, 60170394, 83373044, 20798451, 22279002, 264563, 264488, 264567
3001	88078454 (6001, 6002)	Novel Protein sim. GBank gi 2078470 (AC002073) - Putative gene. Genscan predictions confirmed by EST. splicing: coded for by human cDNAs AA122029 (NID:g1678048), D31562 (NID:g844442), AA158721 (NID:g1733515), R59640 (NID:g830335) and F13082 (NID:g709111) [Homo sapiens]		cathepsin	18108394, 52646842, 56182575, 29331824, 29331825, 29331827, 264910, 33109954, 52644298, 265017, 265019, 264288, 265020, 265021, 52644150, 264692, 35895763, 55810764, 35898423, 56182323, 18108387, 264563, 264564
3002	87718167 (6003, 6004)	Novel Protein sim. GBank gi 3599478 (AF085185) - Myosin IA [Acanthamoeba castellanii]		UNCLASSIFIED	264488, 29331824, 29331825, 29331828, 29331827, 29331828, 264908, 264510, 265009, 21908754, 264682, 264688, 33857023, 264565

3003	86848079 (6005, 6008)	Novel Protein sim. GBank gij1754969 (U30292) - collagen type XIII alpha-1 chain [Mus musculus]	Contains protein domain (PF01391) - Collagen triple helix repeat (20 copies)	collagen	264512, 264593, 264564, 264567, 264488
3004	88066876 (6007, 6008)	Novel Protein sim. GBank gij2224629(dj)BAA20802] - (U002342) KIAA0344 [Homo sapiens]	Contains protein domain (PF01360) - Monooxygenase	oxygenase	26331830, 21906769, 264691, 33657109, 263972, 18108385
3005	87794843 (6008, 6011)	Novel Protein sim. GBank gij4680859(gb)AAD27718.1(AFI3284) CGI-10 protein [Homo sapiens]	Contains protein domain (PF00023) - Ank repeat	MHC	26331822, 29331824, 29331827, 60433438, 265011, 265018, 21908768, 21908767, 21908768, 265020, 33657023, 33657349, 60170394, 22278002, 264587
3006	87422224 (6011, 6012)	Novel Protein sim. GBank gij3930525 (AF064447) - sex-determination protein homolog Fem1a [Mus musculus]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	transcript factor	264259, 29331822, 264512, 21906754, 265018, 264687, 21908765, 264691, 264555, 264558, 264558, 18108385
3007	90936005 (6013, 6014)	Novel Protein sim. GBank gij2565052 (U80736) - CAGH1a [Homo sapiens]	Contains protein domain (PF00501) - AMP-binding enzyme	synthase	52644507, 52845156, 65274572, 264809, 264512, 265018, 264760, 264448, 264765, 264689, 60170815, 18108374, 20281152, 264638, 52644332
3008	80416249 (6015, 6016)	Novel Protein sim. GBank gij3127193 (AF062389) - kidney-specific protein [Rattus norvegicus]	Contains protein domain (PF01923) - Protein of unknown function	UNCLASSIFIED	264905, 264593, 264766, 264636
3009	91213387 (6017, 6018)	Novel Protein sim. GBank gij4927370(gb)AAD33084.1(AF067972) DNA cytosine methyltransferase 3 alpha [Homo sapiens]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	transport	52646842, 56182575, 22278995, 22278996, 264259, 29331825, 29331826, 29331827, 29331828, 35698052, 264508, 264509, 264907, 58182435, 264511, 265007, 264512, 265008, 264757, 264758, 55812038, 264759, 33109954, 21908754, 265010, 265011, 264600, 265017, 265018, 265019, 264760, 18108351, 264288, 264389, 21908784, 21906765, 21906767, 55811957, 265020, 265021, 264691, 18108368, 27486282, 20281149, 18108370, 55811578, 264637, 264558, 264557, 18108381, 264558, 58182323, 264559, 18108385, 18108388, 22278002, 264486
3010	95317217 (6019, 6020)	Novel Protein sim. GBank gij4927370(gb)AAD33084.1(AF067972) DNA cytosine methyltransferase 3 alpha [Homo sapiens]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	transport	264686, 264687, 21906767, 21908768, 55811957, 22278995, 35695917, 22278996, 22278997, 265020, 265021, 60170815, 264692, 33657023, 29331822, 264693, 18108364, 29331824, 33657109, 60432289, 29331827, 27486281, 29331828, 264508, 264909, 55811578, 35695955, 265008, 264558, 60433438, 83373044, 18108387, 65274727, 60432113, 265017, 22278000, 265019, 264584, 264682, 264784
3011	84323597 (6021, 6022)	Novel Protein sim. GBank gij5052318(gb)AAD36501.1(AF11883) citrin; adult-onset type II citrullinemia protein [Homo sapiens]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	transport	35698052, 56182435, 264758, 21908754, 265018, 264760, 264762, 18108351, 264682, 264448, 21908766, 65274620, 18108374, 264482, 264564
3012	87753087 (6023, 6024)		UNCLASSIFIED	UNCLASSIFIED	263972

3013	91238789 (6025, 6028)	Novel Protein sim. GBank gij3702286 (AC005787) - R33374_1 [Homo sapiens]	Contains protein domain (PF00400) - WD domain, G-beta repeat	transcript factor	264488, 263994, 35690286, 22278997, 264259, 29331824, 60424269, 68714117, 3569052, 264905, 264908, 264907, 264908, 264909, 56182435, 264511, 264512, 264910, 264591, 264592, 264593, 264594, 33637402, 60433438, 264595, 264596, 55812038, 264758, 33109954, 21908754, 265010, 265018, 264604, 264760, 264682, 264683, 264764, 264369, 264288, 264765, 264768, 264686, 264768, 264687, 21908767, 35695917, 265020, 33657023, 264692, 264693, 33657109, 264628, 264629, 55811576, 35698423, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264637, 264638, 264639, 83373044, 264563, 264585, 264586, 264567
3014	78877263 (6027, 6028)	Novel Protein sim. GBank gij3878374 [emb CAA93081] - (Z68879) Similarity to Yeast Chi12p protein (PIR Acc. No. S54453); cDNA EST EMBL:D27950 comes from this gene; cDNA EST EMBL:D27949 comes from this gene; cDNA EST EMBL:D33447 comes from this gene; cDNA EST EMBL:D33318 comes from...		ATPase associated	264760
3015	86995466 (6029, 6030)				22278995, 22278998, 22278997, 264259, 29331824, 29331828, 264908, 265007, 265008, 264910, 265011, 265017, 265019, 264691, 33657109, 18108370, 35695855, 264556, 264584
3016	87759945 (6031, 6032)	Novel Protein sim. GBank gij168819 [sp P41733] CC91_YEAST - CELL DIVISION CONTROL PROTEIN 91		UNCLASSIFIED	52644507, 52646842, 56994075, 52645080, 29331822, 29331824, 35696052, 33658970, 52644045, 264596, 33657084, 265017, 265019, 52644229, 21906767, 35695917, 52644150, 33657023, 33657109, 27486261, 27486282, 27486284, 33657349, 27486285, 35695763, 35695855, 87168518
3017	85011154 (6033, 6034)	Novel Protein sim. GBank gij4589658 [dbj BAA76851.1] - (AB023224) KIAA1007 protein [Homo sapiens]			264488, 18108397, 22278996, 35696286, 22278999, 264259, 29331822, 60432289, 264908, 29331830, 264909, 58182435, 265006, 265007, 265008, 265009, 264591, 60433356, 60433438, 52646317, 21908754, 55811388, 265010, 265011, 87168559, 265017, 265018, 265019, 264288, 264687, 21908765, 21908766, 21908767, 21908769, 265020, 265022, 65274620, 52645129, 33657109, 33657182, 18108370, 263972, 18108374, 264631, 52644332, 83373044, 18108385, 18108388, 56526486, 87168518, 264404, 60432113, 22279000, 264567

3018	11073891 (6035, 6036)	Novel Protein sim. GBank gij3219332 (AC004020) - Unknown gene product (Homo sapiens)			264558 264569, 52644507, 18108394, 65274572, 56182575, 22278994, 22278995, 56994075, 22278996, 22278999, 264259, 29331822, 29331824, 60432289, 29331827, 264908, 56182435, 265009, 265009, 60432229, 264593, 60433356, 55812038, 21908754, 37188474, 285011, 87168599, 265017, 265018, 265019, 264681, 18108351, 264448, 264682, 264683, 18108354, 264685, 264687, 264689, 21908766, 21908768, 21908769, 52644150, 264690, 264691, 33657023, 264692, 264693, 33657109, 52645129, 33657349, 264629, 65274791, 264634, 52644332, 56182323, 18108385, 67168518, 22278900, 22278902, 284583
3019	94148231 (6037, 6038)			oncogene	
3020	94318251 (6038, 6040)	Novel Protein sim. GBank gij3414809 (AF061529) - rjs [Mus musculus]	Contains protein domain (PF00415) - Regulator of chromosome condensation (RCC1)	ATPase-associated	264488, 263994, 35696288, 264259, 264508, 264905, 264509, 264908, 264907, 264908, 264909, 264510, 264910, 60174639, 264600, 264603, 264760, 264762, 264682, 264783, 264764, 264288, 264389, 264786, 264687, 264688, 264769, 59811957, 35695917, 33657023, 264628, 35698423, 35695855, 264630, 264632, 264634, 264635, 264636, 264637, 264556, 264557, 264638, 264639, 83373044, 18108385, 264564, 264587, 264486
3021	80478512 (6041, 6042)	Novel Protein sim. GBank gij3880889[emb]CAB09005] - (Z95559) cDNA EST yk238d4.5 comes from this gene; cDNA EST EMBL:C13455 comes from this gene; cDNA EST yk329g8.5 comes from this gene; cDNA EST CEMSH45R comes from this gene [Caenorhabditis elegans]			264769, 264629, 264482
3022	87718500 (6043, 6044)			UNCLASSIFIED	284259, 29331826, 29331828, 264288, 264566
3023	85305484 (6045, 6046)	Novel Protein sim. GBank gij418592[sp]P32323JAGA1_YEAST - A-AGGLUTININ ATTACHMENT SUBUNIT PRECURSOR	Contains protein domain (PF00614) - Phospholipase D. Active site motif	UNCLASSIFIED	284488, 22278995, 35696288, 22278997, 29331826, 35696052, 264907, 29331830, 52644045, 56182435, 60432229, 264592, 60433356, 60433438, 264689, 21908787, 55811957, 35695917, 2649021, 18108376, 263978, 264635, 264558, 22278900
3024	86675305 (6047, 6048)			UNCLASSIFIED	60432049, 264760, 21908769, 55811957, 35695917, 264690, 264555, 264559, 264593, 55811576
3025	65706829 (6049, 6050)	Novel Protein sim. GBank gij295671 (L11275) - selected as a weak suppressor of a mutant of the subunit AC40 of DNA dependant RNA polymerase I and III [Saccharomyces cerevisiae]			

3026	87643862 (6051, 6052)	Novel Protein sim. GBank gij3024052jsp P97924 KARI_RAT - KALIRIN (PAM COOH-TERMINAL INTERACTOR PROTEIN 10) (P-CIP10)		UNCLASSIFIED	22278996, 22278997, 264490, 29331825, 264111, 265007, 60170831, 265010, 87168559, 265019, 21906765, 29148627, 263967, 20281149, 20281069, 263975, 263977, 20281071, 58528486, 22279000
3027	94844563 (6053, 6054)	Novel Protein sim. GBank gij4929847 gb AAD34084.1 AF15184 - (AF151847) CGI-89 protein [Homo sapiens]	Contains protein domain (PF01529) - DHC zinc finger domain	UNCLASSIFIED	18108394, 22278995, 22278996, 35696286, 22278997, 22278998, 60432049, 264259, 29331822, 29331824, 66714117, 29331825, 60432289, 29331826, 29331827, 29331828, 35696052, 264905, 264907, 29331830, 264908, 264909, 264510, 265007, 265008, 265009, 264910, 33657402, 264586, 21906754, 265010, 265011, 87168559, 264600, 265018, 18108351, 264682, 264683, 264764, 264288, 264685, 264687, 264769, 264689, 21906765, 21906768, 21906767, 21906768, 21906769, 29148629, 35695917, 265020, 265021, 265022, 52644150, 264692, 33657023, 264693, 52645129, 33657109, 27486281, 18108374, 55811576, 35696423, 65274791, 264638, 264556, 264557, 264638, 60170394, 264639, 264558, 83373044, 18108385, 58528486, 22279000, 22279002
3028	94231897 (6055, 6056)	Novel Protein sim. GBank gij308052 emb CAA18650] - (AL022598) hypothetical protein [Schizosaccharomyces pombe]	Contains protein domain (PF00400) - WD domain, G-beta repeat	UNCLASSIFIED	22278995, 22278996, 22278997, 22278999, 264259, 29331824, 29331827, 35696052, 264908, 265007, 265008, 265009, 60170831, 21906754, 265011, 87168559, 265018, 284762, 264683, 264765, 264689, 21906765, 21906768, 21906769, 29148629, 35695917, 265021, 265022, 33657109, 27486265, 264628, 264629, 18108374, 35696423, 35695855, 264638, 60170394, 22279000, 22279002, 264482, 264564
3029	87619284 (6057, 6058)			UNCLASSIFIED	22278997, 22278999, 29331827, 264905, 264509, 264909, 264510, 264511, 264512, 87168474, 265019, 18108351, 21906768, 264534, 264690, 264693, 263989, 18108370, 264558, 22279000, 22279002, 264482
3030	87544828 (6059, 6060)	Novel Protein sim. GBank gij3757726 emb CAA18782] - (AL022727) dJ8019.1 (olfactory receptor-like protein [hs6M1-1]) [Homo sapiens]	Contains protein domain (PF00001) - 7 transmembrane receptor (rhodopsin family)	lm7	
3031	91677953 (6061, 6062)	Novel Protein sim. GBank gij4530587 gb AAD22105.1] - (AF132000) TADA1 protein [Homo sapiens]		UNCLASSIFIED	22278995, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 29331826, 29331827, 29331828, 35696052, 264908, 265018, 21906765, 21906766, 21906767, 21906768, 265021, 263974, 18108374, 264558, 58528486, 22279000, 22279002

3032	94130124 (6063, 6084)	Novel Protein sim. GBank gi 1019951 (U37429) - similar to M. musculus MERS and other AHPC/TSA proteins [Caenorhabditis elegans]	Contains protein domain (PF00534) - synthase Glycosyl transferases group 1	22278996, 35696286, 264259, 29331824, 29331828, 284807, 29331830, 264758, 33109954, 87168474, 87168559, 285019, 264288, 21908769, 285021, 264683, 35696423, 35695855, 264636, 56182323, 83373044, 87168518
3033	95308321 (6065, 6086)	Novel Protein sim. GBank gi 5031573 ref NP_005712.1 PACTR - ARP3 (actin-related protein 3, yeast) homolog	Contains protein domain (PF00022) - struct Actin	35696286, 264259, 29331826, 35696052, 284508, 264905, 264806, 264907, 284908, 264909, 265008, 264591, 21908754, 265010, 265019, 264681, 264369, 264768, 21908764, 21908768, 35695917, 33657023, 264628, 35695855, 264632, 264635, 264639, 264482, 264563
3034	80415373 (6067, 6088)		UNCLASSIFIED	284908, 284907, 264510, 264592, 265010, 284782, 264766, 264637, 264638, 264486
3035	91220892 (6069, 6070)	Novel Protein sim. GBank gi 3738207 emb CAA21282 (AL031853) conserved ATP-GTP binding protein [Schizosaccharomyces pombe]	UNCLASSIFIED	264636
3036	91718323 (6071, 6072)	Novel Protein sim. GBank gi 728837 sp P39194 ALU7_HUMAN - III ALU SUBFAMILY SQ WARNING ENTRY III	kinase	264807, 33657402, 265021
3037	95307434 (6073, 6074)	Novel Protein sim. GBank gi 4406590 gb AAD20040 (AF131766) Similar to Ena-VASP like protein [Homo sapiens]		285017
3038	95421807 (6075, 6076)	Novel Protein sim. GBank gi 5380093 gb AAD42865.1 AF15509 - (AF15509) NY-REN 18 antigen [Homo sapiens]	Contains protein domain (PF00627) - UNCLASSIFIED UBA domain	22278998, 22278997, 264259, 284905, 285007, 265009, 60433356, 21908754, 265016, 265019, 18108351, 264687, 21908765, 265020, 265021, 65274820, 27486282, 264638, 56182323, 18108385, 22278000
3039	87332257 (6077, 6078)	Novel Protein sim. GBank gi 475128 emb CAB42094.1 (AJ238717) ZRP protein [Rattus norvegicus]	UNCLASSIFIED	35696286, 29331828, 264109, 264110, 264511, 265007, 21908754, 265011, 264681, 264683, 264687, 21908768, 264691, 18108370, 263972, 264629, 18108374, 263977, 35696423, 264584, 18108391
3040	80933517 (6079, 6080)	Novel Protein sim. GBank gi 4884278 emb CAB43247.1 (AL050037) hypothetical protein [Homo sapiens]		264682, 264558, 18108382, 18108385, 284587
3041	88312357 (6081, 6082)	Novel Protein sim. GBank gi 3876073 emb CAB04122.1 (Z81505) similar to Zinc finger, C3HC4 type (RING finger); cDNA EST EMBL:D28025 comes from this gene; cDNA EST EMBL:D28024 comes from this gene; cDNA EST EMBL:D33210 comes from this gene; cDNA EST EMBL:D33441 comes from this ...	UNCLASSIFIED	56994075, 22278997, 22278998, 29331827, 33656970, 33108954, 21908754, 87168559, 264600, 264683, 21908765, 21908768, 22278002
3042	85749402 (6083, 6084)	Novel Protein sim. GBank gi 790236 (U21156) - sarcolemmal associated protein-2 [Oryctolagus cuniculus]	glycoprotein	264636

3043	87773026 (6085, 6086)	Novel Protein sim. GBank gi 854065 emb CAA58337 - (X83413) U88 [Human herpesvirus 6]		UNCLASSIFIED	35696286, 60424268, 35696052, 264508, 264905, 68712502, 56182435, 55811386, 52644298, 55811150, 35695917, 60170615, 33657109, 18108374, 284834, 60431850
3044	87646182 (6087, 6088)	Novel Protein sim. GBank gi 4104922 (AF042276) - c251 homolog [Pseudomonas putida]	Contains protein domain (PF01208) - ubiE/COQ5 methyltransferase family	glycoprotein	22278998, 22278998, 22278998, 29331824, 56182435, 264511, 265007, 60170831, 60432229, 60433356, 33109954, 18108351, 264288, 35695917, 18108388, 18108370, 60170394
3045	94127598 (6089, 6090)	Novel Protein sim. GBank gi 4589860 dbj BAA76859.1 - (AB023232) KIAA1015 protein [Homo sapiens]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	dna_mg_bind	264488, 264259, 35696052, 264508, 264905, 264509, 264906, 284907, 264909, 264511, 265006, 264591, 264593, 33109954, 284604, 284764, 264683, 284288, 264768, 264768, 21906765, 21906768, 55811957, 35695917, 27486262, 18108370, 264628, 18108374, 35695855, 264630, 264632, 264635, 264563, 264564, 264566
3046	88098247 (6091, 6092)			UNCLASSIFIED	22278999, 29331822, 29331824, 29331825, 29331826, 60432289, 29331827, 29331828, 264906, 52846317, 55811957, 60432113, 22278000, 22278002, 264482, 264584
3047	95089924 (6093, 6094)			UNCLASSIFIED	264488, 22278998, 22278997, 22278999, 29331824, 29331825, 56182435, 264511, 265008, 285009, 265011, 265017, 284768, 21906768, 21906769, 35695917, 52844150, 33657349, 65274791, 35695855, 264555, 60432113, 22279000, 264568
3048	87629419 (6095, 6096)	Novel Protein sim. GBank gi 4588034 gb AAD25962.1 AF09287 - (AF092878) zinc RING finger protein SAG [Homo sapiens]	Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger)	UNCLASSIFIED	264102, 29148784
3049	88229955 (6097, 6098)	Novel Protein sim. GBank gi 5454158 ref NP_006286.1 pVARS - valyl-IRNA synthetase 1	Contains protein domain (PF01406) - tRNA synthetases class I (C)	UNCLASSIFIED	22278997, 29331826, 264907, 264758, 87168559, 265018, 264448, 21908768, 265020, 33657109, 35695855, 60432113, 22279000
3050	87643679 (6099, 6100)	Novel Protein sim. GBank gi 4589842 dbj BAA76843.1 - (AB023218) KIAA0999 protein [Homo sapiens]	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	kinase	264259, 29331825, 264909, 265007, 264512, 265019, 264288, 21908768, 265020, 264893, 18108385, 56526466, 87168518, 22279002, 264566
3051	87750599 (6101, 6102)				22278997, 264595, 265019, 264288, 264693, 87168518
3052	57108030 (6103, 6104)	Novel Protein sim. GBank gi 117528 sp P14755 CRYL_RABIT - LAMBDA-CRYSTALLIN		dehydrogenase	264534

3053	95350373 (6105, 6106)	Novel Protein sim. GBank gll3947613[emb]CAA19465.1] - (AL023826) cDNA EST EMBL:M89008 comes from this gene; cDNA EST yk282d3.5 comes from this gene [Caenorhabditis elegans]		UNCLASSIFIED	65274572, 56181666, 22278995, 35696286, 22278998, 264259, 60432289, 265008, 265009, 60433438, 21908754, 265010, 87108559, 264603, 265018, 265019, 264763, 264764, 264288, 21908765, 21908768, 21908769, 35695917, 18108374, 35696423, 264638, 56182323, 22279000, 264563
3054	86943510 (6107, 6108)	Novel Protein sim. GBank gll1076211[pi]SS0755 - hypothetical protein VSP-3 - Chlamydomonas reinhardtii		UNCLASSIFIED	25696288, 35696052, 28331830, 264808, 264908, 264512, 264910, 265017, 264604, 264766, 265020, 33657109, 264628, 35695955, 264638, 264564, 264566, 264486, 60424178, 65274572, 56182575, 35696286, 22278998, 22278999, 60432048, 264259, 60424269, 60432289, 35696052, 56182435, 265008, 265009, 60170831, 60432229, 60431735, 60433356, 264594, 60433438, 21908754, 55811386, 265011, 87168559, 265019, 18108351, 264683, 264288, 264369, 264689, 21908768, 55811957, 35695917, 60170815, 33657023, 65274620, 33657109, 35695763, 60431528, 18108374, 55810764, 55811576, 35696423, 65274791, 264638, 60431850, 18108381, 56182323, 60170394, 18108385, 60432113, 264564, 264565, 264566
3055	95350537 (6109, 6110)	Novel Protein sim. GBank gll4680655[gb]AAD27717.1[AF13294 - (AF132842) CGI-08 protein (Homo sapiens)]		transport	264488, 264569, 18108394, 52646842, 22278997, 22278998, 22278999, 264259, 66714117, 29331826, 29331827, 35696052, 264508, 264509, 264905, 264906, 264807, 264908, 264909, 265006, 264512, 265007, 265008, 265009, 264910, 33657402, 55812038, 264596, 264758, 265010, 265011, 265017, 265018, 264780, 18108331, 264762, 264763, 264764, 264288, 264768, 264687, 18108357, 264768, 264769, 264689, 21908765, 21908766, 21908767, 21908768, 21908769, 35695917, 265020, 265021, 264691, 264693, 33657109, 18108370, 264628, 264629, 18108374, 55811576, 35696423, 35695855, 264630, 264631, 264632, 264634, 264635, 264638, 264637, 264638, 18108381, 83373044, 18108385, 22279000, 22279002, 264563, 264564, 264565, 264566, 264486, 264567
3056	91661638 (6111, 6112)	Novel Protein sim. GBank gll728637[sp]P39184[ALU7 - HUMAN - III] ALU SUBFAMILY SQ WARNING ENTRY IIII		glycoprotein	

3057	95412746 (6113, 6114)	Novel Protein sim. GBank gi 3878119 emb CAA8860 - (Z49088) similar to GTP-binding protein; cDNA EST EMBL:M89111 comes from this gene; cDNA EST EMBL:D27709 comes from this gene; cDNA EST EMBL:D27708 comes from this gene; cDNA EST EMBL:D73788 comes from this gene; cDNA EST A353...			struct	264508, 264905, 264907, 264908, 264909, 264510, 264512, 264910, 264592, 264594, 264767, 18108374, 264635, 264555, 264637, 264639, 264583, 264564, 264565, 264486
3058	79646228 (6115, 6116)	Novel Protein sim. GBank			UNCLASSIFIED	264693
3059	87628425 (6117, 6118)	gi 4588034 gb AAD25982.1 AF092878) zinc RING finger protein SAG [Homo sapiens]	Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger)		UNCLASSIFIED	22278995, 22278996, 22278997, 22278998, 22278999, 264490, 264259, 29331824, 29331825, 29331827, 35696032, 29331828, 265007, 60433438, 265017, 265018, 265019, 264681, 264448, 264288, 264768, 21908765, 21908766, 21908767, 21908769, 29148629, 29148784, 265022, 52844150, 18108370, 264636, 18108385, 264563, 264567
3060	79346691 (6119, 6120)				UNCLASSIFIED	264567
3061	87740964 (6121, 6122)				UNCLASSIFIED	264112, 52844296, 21908768, 33657023, 263974, 18108385
3062	87619465 (6123, 6124)	Novel Protein sim. GBank gi 4454680 gb AAD20963) - (AF070857) glutathione S-transferase subunit 13 homolog [Homo sapiens]			transferase	264908, 265008, 18108351, 264568
3063	80078023 (6125, 6126)	Novel Protein sim. GBank gi 2246532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus]			UNCLASSIFIED	18108359, 264558
3064	91241526 (6127, 6128)	Novel Protein sim. GBank gi 4240315 dbj BAA74936.1) - (AB020720) KIAA0913 protein [Homo sapiens]	Contains protein domain (PF00403) - Heavy-metal-associated domain		UNCLASSIFIED	52846385, 52846842, 65274572, 56182575, 56181686, 22278995, 22278996, 22278997, 22278998, 264259, 60432049, 29331824, 66714117, 264508, 264907, 264908, 56182435, 265009, 60432229, 60433438, 55812038, 52644286, 265018, 264682, 264288, 264686, 264768, 264687, 52644229, 264689, 21908768, 264691, 264692, 264693, 18108370, 18108377, 55811576, 264638, 56182323, 264558, 264639, 18108385, 22279000, 22279002
3065	91639201 (6129, 6130)	Novel Protein sim. GBank gi 5656743 gb AAD45960.1 AC005067) Supported by Human EST H08032.1 (NID:g872854), mouse EST AA870042.1 (NID:g2985487), and gencscan [Homo sapiens]			UNCLASSIFIED	22278996, 22278998, 264093, 264094, 264095, 29331824, 60424269, 66714117, 264100, 264907, 265007, 264591, 60432229, 264593, 265011, 265019, 18108351, 264768, 264767, 21908765, 21908768, 264693, 20281069, 22279000, 22279002, 264482, 264566, 264567

3066	91224437 (6131, 6132)	Novel Protein sim. GBank gl 4884268 emb CAB3245.1) - (AL050028) hypothetical protein [Homo sapiens]		UNCLASSIFIED	18108397, 22278995, 56994075, 22278998, 264905, 66712502, 265008, 264512, 264910, 264758, 60174639, 264780, 18108351, 264784, 264683, 18108359, 264692, 18108364, 18108368, 18108370, 18108377, 18108379, 60170394, 264587
3067	95422551 (6133, 6134)	Novel Protein sim. GBank gl 4889258 gb AAD27832.1 AF12185 - (AF121859) sorting nexin 9 [Homo sapiens]	Contains protein domain (PF00787) - struct PX domain		264488, 264489, 35696286, 22278998, 56994075, 264259, 29331822, 29331825, 35698052, 29331828, 264508, 264905, 264509, 264908, 264907, 264908, 264909, 264112, 264510, 264511, 264512, 265008, 265009, 264910, 264591, 264592, 264593, 264594, 264757, 264595, 264598, 264758, 265010, 265011, 87168559, 264801, 264802, 264603, 264604, 264605, 265019, 264760, 264762, 264448, 264763, 264764, 264288, 264389, 264766, 264768, 264687, 264769, 264689, 21908765, 21908767, 21908768, 35695917, 265020, 265021, 264534, 52844150, 264691, 33857023, 264893, 264628, 60431528, 263977, 35695855, 264630, 264631, 264634, 264635, 264636, 264637, 264638, 264639, 83373044, 56328488, 87168518, 22278998, 22279002, 264563, 264483, 264564, 264565, 264566, 264567, 264486
3068	85380851 (6135, 6136)				264112
3069	95412753 (6137, 6138)	Novel Protein sim. GBank gl 3878119 emb CAA88860) - (Z49068) similar to GTP-binding protein; cDNA EST EMBL:M89111 comes from this gene; cDNA EST EMBL:D27709 comes from this gene; cDNA EST EMBL:D27708 comes from this gene; cDNA EST EMBL:D73788 comes from this gene; cDNA EST yk353...	Contains protein domain (PF01928) - struct GTPase of unknown function		22278998, 56994075, 22278998, 22278999, 264259, 264107, 264905, 29331830, 52844045, 264110, 60170831, 264592, 264594, 33657402, 21908754, 33109954, 87168474, 87168559, 265017, 264448, 264764, 264683, 264766, 52644229, 21908765, 21908768, 21908768, 21906769, 60170815, 33657023, 18108370, 18108376, 264634, 264557, 60170394, 56182323, 18108385, 87168518, 22279000, 264482

3070	94319173 (6139, 6140)	Novel Protein sim. GBank gll3877788jemb[CAB05527] - (Z83110) cDNA EST yk472b5.3 comes from this gene; cDNA EST yk474a7.3 comes from this gene; cDNA EST yk472b5.5 comes from this gene; cDNA EST yk468c10.3 comes from this gene; cDNA EST yk468c10.5 comes from this gene; cDNA EST EM...		synthase	284488, 22278994, 22278995, 22278996, 58994075, 22278997, 22278999, 284259, 29331822, 29147620, 29331824, 68714117, 29331826, 29146498, 29146499, 68712502, 29331830, 52844045, 58182435, 284511, 265007, 284512, 284910, 60170831, 284592, 284758, 33108954, 21908754, 87188474, 265019, 18108351, 284448, 284683, 284288, 52844229, 284689, 21908765, 21908768, 21908767, 21908769, 35895917, 265020, 265021, 60170815, 52844150, 284691, 33857023, 27486281, 27486284, 284628, 18108370, 18108377, 55811578, 35895855, 284634, 284635, 18108381, 60170394, 58182323, 284558, 83373044, 18108385, 18108387, 56526488, 264404, 284583, 284586
3071	94325573 (6141, 6142)	Novel Protein sim. GBank gll4502425jefjNP_001709.1pBMP6 - bone morphogenetic protein 6 precursor	Contains protein domain (PF00085) - Igt Thioedoxin		284488, 65274572, 18108398, 22278996, 35898288, 22278997, 22278998, 22278999, 284259, 29331822, 66714117, 29331826, 35898052, 29331828, 29146498, 29146499, 284907, 284908, 29331830, 284909, 52844045, 58182435, 285008, 285007, 284512, 265008, 265009, 60170831, 60432228, 284592, 60433358, 33857402, 60433438, 33109954, 52844296, 87188474, 285010, 285017, 284681, 284288, 284685, 284766, 284687, 284769, 284689, 21908765, 21908768, 21908767, 21908768, 21908769, 35895917, 265020, 265021, 265022, 60170815, 52844150, 284690, 284691, 284692, 33857023, 284693, 33857109, 283971, 18108374, 18108377, 35898423, 55811578, 65274791, 35895855, 284630, 284635, 284636, 284557, 60170394, 83373044, 60432113, 22279000, 22279002, 284583, 284584, 284585, 284586, 284587, 56182575, 29331822, 29331824, 29331825, 29146498, 284908, 52844045, 58182435, 265009, 60433438, 55812038, 18108351, 284683, 284369, 52844229, 52844150, 33857023, 284693, 33857109, 18108374, 55811578, 65274791, 284555, 58182323, 60432113, 284584
3072	95115892 (6143, 6144)	Novel Protein sim. GBank gll1263289 (U47856) - fibrin-4 [Araneus diadematus]		transcriptfactor	

3073	86147248 (6145, 6146)	Novel Protein sim. GBank gi1134840 sp P22523 CORB_HUMAN - CORNIFIN B (SMALL PROLINE-RICH PROTEIN IB) (SPR-IB) (14.9 KD PANCORNULIN)		UNCLASSIFIED	264769
3074	88089351 (6147, 6148)	Novel Protein sim. GBank gi134840 sp P22523 CORB_HUMAN - CORNIFIN B (SMALL PROLINE-RICH PROTEIN IB) (SPR-IB) (14.9 KD PANCORNULIN)		UNCLASSIFIED	264488, 265019, 264448, 264288, 21908767, 264893, 18108388, 18108370, 18108374, 264587
3075	88095752 (6149, 6150)	Novel Protein sim. GBank gi134840 sp P22523 CORB_HUMAN - CORNIFIN B (SMALL PROLINE-RICH PROTEIN IB) (SPR-IB) (14.9 KD PANCORNULIN)	Contains protein domain (PF00023) - Ank repeat	homeobox	264509, 264907, 264689, 264693, 56526486
3076	87819219 (6151, 6152)	Novel Protein sim. GBank gi134840 sp P22523 CORB_HUMAN - CORNIFIN B (SMALL PROLINE-RICH PROTEIN IB) (SPR-IB) (14.9 KD PANCORNULIN)		UNCLASSIFIED	18108398, 29331822, 29331827, 60432229, 265017, 264691, 264693
3077	88734277 (6153, 6154)	Novel Protein sim. GBank gi13023958 sp Q00808 HET1_PODAN - VEGETATIBLE INCOMPATIBILITY PROTEIN HET-E-1	Contains protein domain (PF00400) - WD domain, G-beta repeat	kinase	65274572, 35696052, 264511, 60170831, 87168474, 264389, 35695917, 33657182, 27486264, 33657349, 35695763, 35695855, 264639
3078	88089355 (6155, 6156)	Novel Protein sim. GBank gi13023958 sp Q00808 HET1_PODAN - VEGETATIBLE INCOMPATIBILITY PROTEIN HET-E-1			22278002
3079	87821893 (6157, 6158)	Novel Protein sim. GBank gi13875410 emb CAB02876 - (Z81052) Similarity to Yeast ABC1P protein (SW:ABC1_YEAST); cDNA EST yk229g8.3 comes from this gene; cDNA EST yk229g8.5 comes from this gene (Caenorhabditis elegans)		transport	29331824, 29331826, 264758, 55811388, 265017, 55811150, 52644228, 21908788, 265020, 265021, 264693, 18108378, 264631, 52644332, 22278002
3080	93288274 (6159, 6160)	Novel Protein sim. GBank gi15257221 gb AAD41265.1 - (AF117887) protein arginine methyltransferase [Mus musculus]		interferon	264488, 52644507, 22278998, 22278998, 264490, 264259, 29331824, 86714117, 29331825, 29331828, 29331827, 29331828, 29148489, 264508, 264905, 264828, 52644045, 56182435, 265008, 264591, 264596, 21908754, 60174639, 265010, 264882, 264448, 264763, 264683, 264784, 264288, 264685, 264769, 264688, 264689, 21908765, 21908767, 21908769, 55811857, 35695917, 265020, 60170815, 52644150, 264692, 33657023, 264693, 65274820, 33657109, 27486261, 35695763, 264628, 18108370, 65274791, 264558, 56182323, 60170394, 264482, 264585, 264484
3081	88084864 (6161, 6162)	Novel Protein sim. GBank gi1728831 sp P39188 ALU1_HUMAN - IIII ALU SUBFAMILY J WARNING ENTRY IIII		UNCLASSIFIED	18108398, 264509, 264905, 264908, 264907, 264908, 264909, 264510, 264511, 265009, 264910, 264595, 264758, 265011, 265018, 264760, 264781, 264763, 264784, 18108354, 264685, 264766, 264628, 264629, 264630, 264631, 264632, 264634, 264635, 264555, 264638, 18108382, 18108385, 264583, 264565, 264568
3082	80310121 (6163, 6164)				264784, 55811857, 264555, 264584

3083	88095756 (6165, 6166)	Novel Protein sim. GBank gj1868241 (U29488) - C56C10.3 gene product [Caenorhabditis elegans]		UNCLASSIFIED	284488, 284259, 29331824, 284106, 265008, 284591, 284592, 21906754, 284288, 264767, 21906768, 21906769, 29148784, 264691, 264632, 22279000
3084	87448568 (6167, 6168)	Novel Protein sim. GBank gj1476774 (pir1jA37475 - probable structural component p38 - borna disease virus			22278995, 60432289, 35696052, 264905, 284906, 264907, 264908, 264909, 265006, 265007, 264910, 284593, 264595, 264758, 264368, 264288, 264766, 35695917, 265020, 18108374, 35696423, 264631, 264556, 284585, 264566, 264567, 264486
3085	87785781 (6169, 6170)	Novel Protein sim. GBank gj12585057 (U80741) - CAGH44 [Homo sapiens]		UNCLASSIFIED	265011, 264681
3086	87769942 (6171, 6172)	Novel Protein sim. GBank gj13894189 (AC005662) - hypothetical protein [Arabidopsis thaliana]		UNCLASSIFIED	22278998, 264092, 264258, 29331822, 29331825, 264108, 264112, 18108351, 264687, 263987, 263974, 55810764, 263981, 18108385, 284487
3087	87462888 (6173, 6174)				52846365, 56994075, 22278997, 22278998, 29331824, 29331825, 35696052, 60433438, 33109954, 21906754, 52846317, 285017, 264682, 264369, 264684, 21908767, 21906768, 265020, 264691, 33857023, 33657109, 52645129, 33657182, 27488262, 35695855, 87188518
3088	91224441 (6175, 6176)	Novel Protein sim. GBank gj13353304 (AF001549) - Unknown gene product [Homo sapiens]		UNCLASSIFIED	264591
3089	95361242 (6177, 6178)	Novel Protein sim. GBank gj14689146 (gb AAD27782.1 AF07704 - (AF077049) lambda-crystallin [Homo sapiens]	Contains protein domain (PF00725) - dehydrogenase 3-hydroxyacyl-CoA dehydrogenase		18108397, 65274572, 56182575, 56181686, 56994075, 35696286, 22278997, 22278998, 264259, 29331824, 29331825, 29331826, 29331828, 284907, 29331830, 284909, 58182435, 264510, 265007, 60170831, 60432229, 21906754, 55811386, 285017, 265016, 265019, 264760, 55811150, 264288, 264766, 56181562, 21906765, 21906766, 21906767, 21906768, 265021, 60170815, 27486262, 18108370, 60431528, 35696423, 264558, 264559, 60432113, 264486

3090	95342371 (6179, 6180)	Novel Protein sim. GBank g11354050 (U47024) - MEM3 [Mus musculus]		UNCLASSIFIED	60424179, 52645156, 65274572, 58182575, 56181886, 22278995, 35898288, 58994075, 22278998, 22278998, 22278999, 284259, 29331822, 56182181, 29331824, 29331825, 29331826, 29331827, 29331828, 35896052, 33856970, 264908, 264908, 52644045, 264828, 265008, 265007, 265008, 60170831, 60432229, 60433356, 33857402, 55812038, 264758, 21908754, 33109954, 52648317, 55811386, 52644296, 87188474, 265011, 87168559, 265017, 265018, 265019, 55811150, 18108351, 264881, 264448, 264288, 264368, 18108357, 264768, 52844228, 58181582, 21908764, 21906765, 21908766, 21908767, 21908768, 21908769, 35895917, 265020, 265022, 60170815, 264890, 52644150, 264691, 33857023, 18108365, 65274820, 33857109, 18108368, 33857182, 27486261, 27486265, 35895763, 18108374, 18108376, 55810764, 35898423, 55811578, 65274791, 35895855, 264557, 58182323, 83373044, 18108387, 18108388, 87188518, 22279000, 22278002, 264583, 264482
3091	95317424 (6181, 6182)	Novel Protein sim. GBank g13873932[emb]CAB01859] - (Z79598) Similarity to Bovine aspartyl beta hydroxylase (TR:G162694); cDNA EST EMBL:D27916 comes from this gene; cDNA EST EMBL:D27915 comes from this gene; cDNA EST EMBL:D64881 comes from this gene; cDNA EST EMBL:D68139 comes f...		UNCLASSIFIED	35898286, 29331822, 35896052, 264508, 264509, 264905, 264908, 264908, 264909, 264510, 264758, 265010, 265011, 264883, 264885, 264768, 264768, 264769, 264693, 264628, 35898423, 35895855, 264632, 264635, 264839, 264482, 264583, 264486

3082	95314502 (6183, 6184)	Novel Protein sim. GBank gi 1710756 sp P15880 RS2_HUMAN - 40S RIBOSOMAL PROTEIN S2 (S4) (LLREP3 PROTEIN)	Contains protein domain (PF00333) - ribosomal prot Ribosomal protein S5	264488, 60424179, 18108398, 22278995, 59994075, 22278998, 35698286, 22278997, 22278998, 60432049, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 35696052, 29331828, 29146498, 29146499, 264508, 264509, 264905, 264906, 264907, 29331830, 264908, 264909, 264113, 264510, 264511, 265008, 264512, 265007, 265008, 264910, 265009, 60170831, 264591, 264592, 60431735, 264593, 264594, 60433438, 264595, 264758, 21908754, 265010, 265011, 264601, 264602, 265017, 264603, 264604, 265018, 264605, 265019, 264760, 264762, 264681, 18108351, 264763, 264682, 264448, 264764, 264683, 264288, 264389, 264765, 264766, 264686, 264767, 264687, 264768, 264769, 264688, 21908764, 264689, 21908765, 21908766, 21908767, 21908768, 21908769, 29148629, 29148784, 35695917, 265020, 265021, 264534, 60170615, 264690, 264691, 264692, 65274820, 33857109, 27486262, 264628, 264629, 18108374, 263978, 18108377, 35696423, 264630, 264631, 264632, 264634, 264635, 264555, 264636, 264637, 264556, 264638, 264557, 264558, 264639, 60170394, 18108385,
3093	94318457 (6185, 6186)	Novel Protein sim. GBank gi 5002587 emb CAB44347.1 - (Y17454) LSFR1 protein [Homo sapiens]	UNCLASSIFIED	264259, 29331824, 35696052, 264905, 265008, 60432229, 60431735, 264684, 264369, 264288, 264766, 21908767, 35698423, 63373044, 18108385
3094	94316675 (6187, 6188)	Novel Protein sim. GBank gi 400734 sp P31044 PBP_RAT - PHOSPHATIDYLETHANOLAMINE-BINDING PROTEIN (23 KD MORPHINE-BINDING PROTEIN) (P23K)	Contains protein domain (PF01161) - Phosphatidylethanolamine-binding protein	18108398, 264259, 60432289, 29331827, 264511, 264763, 264288, 264767, 265022, 264691, 264693, 65274791, 56182323, 264584, 264565

3095	94848162 (6189, 6190)	Novel Protein sim. GBank g[487759]g[AD31421.1]AF12444 - (AF12444) MAGE tumor antigen D1 [Homo sapiens]	Contains protein domain (PF01454) - MAGE family	UNCLASSIFIED	18108397, 56182575, 22278995, 35896286, 5694075, 22278997, 22278999, 264259, 60432049, 66714117, 29331825, 60432289, 35898052, 33656970, 29148499, 284508, 284905, 264509, 29331830, 264908, 264510, 264511, 284512, 265007, 265008, 265009, 60170831, 264758, 21906754, 85658542, 265010, 265011, 87168559, 265017, 265018, 265019, 264760, 264681, 264682, 264683, 264764, 264368, 264288, 264886, 264788, 264769, 264889, 21908765, 21908766, 21908787, 55811957, 35895917, 265020, 265021, 265022, 52644150, 264691, 264692, 33857023, 264693, 263972, 18108378, 55811576, 35696423, 264952, 60170394, 264639, 83373044, 18108385, 18108387, 65274727, 87168518, 60432113, 264482, 264563, 264564, 264566, 264487, 18108391, 22278995, 22278996, 22278997, 22278999, 29331824, 29331825, 29331828, 29331827, 33856970, 264905, 264908, 265008, 264910, 33857402, 265011, 265017, 265018, 264389, 21908786, 21908787, 21908788, 35695917, 265020, 60170815, 264691, 264692, 264693, 27486261, 27486262, 18108370, 60431528, 264634, 264636, 264639, 22279000, 264566, 264488, 29331822, 29331825, 60432289, 29331826, 35696052, 29331828, 29331830, 264594, 55812038, 33109854, 33657084, 87168474, 87168559, 52644229, 21908765, 21908787, 18108378, 35696423, 52644332, 264638, 60432113, 22279002, 264634, 264637, 264565
3096	87756128 (6191, 6192)	Novel Protein sim. GBank g[3882221]dbj[BAA34470.1] - (ABD18283) KIAA0750 protein [Homo sapiens]	Contains protein domain (PF00307) - struct Calponin homology (CH) domain		
3097	88264895 (6193, 6194)	Novel Protein sim. GBank g[4488288]emb[CAB37981] - (AL022395) dJ273N12.1 (PUTATIVE protein based on EST matches) [Homo sapiens]	Contains protein domain (PF00646) - F-box domain.	UNCLASSIFIED	264488, 29331822, 29331825, 60432289, 29331826, 35696052, 29331828, 29331830, 264594, 55812038, 33109854, 33657084, 87168474, 87168559, 52644229, 21908765, 21908787, 18108378, 35696423, 52644332, 264638, 60432113, 22279002, 264634, 264637, 264565
3098	80258024 (6195, 6196)	Novel Protein sim. GBank g[303603]dbj[BAA02145.1] - (D12621) cytochrome P-450 1B1 [Homo sapiens]		cyto450	264488, 35696286, 29331822, 29331824, 29331825, 29331827, 265007, 265008, 265010, 265011, 265018, 265019, 18108357, 21908766, 265020, 265022, 55811578, 56182323, 22279002, 264563
3099	91243325 (6197, 6198)	Novel Protein sim. GBank g[1083764]pir[B48013] - proline- rich proteoglycan 2 precursor, paroloid - rat		UNCLASSIFIED	29331825, 60432289, 35696052, 264910, 60432228, 264592, 264288, 264693, 263987, 264635
3100	87602421 (6199, 6200)			UNCLASSIFIED	264808, 264693, 264628, 264630, 264632
3101	79602134 (6201, 6202)			UNCLASSIFIED	

3102	91220892 (6203, 6204)	Novel Protein sim. GBank gij5305708 gb AAD41781.1 AF12853 - (AF128536) Cytoplasmic phosphoprotein PACSIN2 [Homo sapiens]	Contains protein domain (PF00018) - SH3 domain	SH3 domain	35696286, 22278996, 22278999, 29331827, 35696052, 264909, 264512, 265008, 60170831, 60433356, 33109954, 18108351, 264684, 264689, 21906767, 60170615, 264692, 33657023, 264638, 22279000, 264482, 264564
3103	90938004 (6205, 6206)	Novel Protein sim. GBank gij464564 sp P35292 RB17_MOUSE - RAS-RELATED PROTEIN RAB-17	UNCLASSIFIED	UNCLASSIFIED	35695917, 264565
3104	67340833 (6207, 6208)	Novel Protein sim. GBank gij5032207 ref NP_005698.1 pTSSC - tumor-suppressing STF cDNA 6	UNCLASSIFIED	UNCLASSIFIED	264259, 264684, 264532, 33657182, 264558
3105	94148503 (6209, 6210)				22278997, 264259, 29331824, 35696052, 29331828, 264508, 264509, 264905, 264906, 264907, 264908, 264511, 264910, 264591, 264594, 264758, 264760, 264681, 264762, 264764, 264288, 264768, 264768, 264687, 264789, 21906768, 21906768, 35695917, 33657023, 264692, 264693, 264628, 264629, 35695855, 264630, 264631, 264632, 264634, 264635, 264637, 264638, 264639, 83373044, 264404, 22278002, 264563, 264565, 264566, 264486, 264567
3106	95361416 (6211, 6212)	Novel Protein sim. GBank gij1938574 (U97190) - B0025.2 gene product [Caenorhabditis elegans]			22278996, 22278997, 22278998, 22278999, 264092, 264093, 264094, 29331822, 264908, 264907, 264908, 52644045, 58182435, 264112, 265008, 265009, 55812038, 265017, 285018, 264683, 264688, 264687, 264788, 52644229, 21906765, 21906768, 21906769, 55811957, 265020, 265022, 264690, 52644150, 264692, 264693, 18108370, 18108377, 55811576, 56182323, 18108385, 18108388, 22279000, 264563
3107	95343272 (6213, 6214)	Novel Protein sim. GBank gij3341441 emb CAA76851 - (Y17794) winged-helix transcription factor [Gallus gallus]			22278995, 22278996, 35696286, 22278997, 22278999, 264091, 264093, 264259, 29331822, 29331825, 29331826, 60432289, 29331827, 29331828, 33656970, 264105, 264512, 265009, 60433358, 60433438, 265011, 265017, 265018, 21906765, 21906768, 21906767, 21906769, 265021, 264691, 33657109, 27486261, 27486265, 18108370, 263972, 18108374, 55811576, 18108385, 56526486, 264482, 264487, 56182435, 264288, 264690, 264564
3108	67340835 (6215, 6216)	Novel Protein sim. GBank gij5032207 ref NP_005698.1 pTSSC - tumor-suppressing STF cDNA 6	UNCLASSIFIED	UNCLASSIFIED	

3109	84318461 (6217, 6218)	Novel Protein sim. GBank gl5002587[emb]CAB44347.1] - (Y17454) LSFRT protein [Homo sapiens]	Contains protein domain (PF00008) - Zinc finger, C2H2 type	stud1	264480, 264808, 265007, 264910, 264593, 264883, 264684, 264687, 21906767, 21908788, 264693, 18108370, 264629, 18108374, 264632, 264638, 22279000
3110	85090718 (6219, 6220)	Novel Protein sim. GBank gl1078211[pir]S50755 - hypothetical protein VSP-3 - Chlamydomonas reinhardtii		UNCLASSIFIED	264488, 65274572, 22278995, 22278897, 60432049, 264259, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 29331828, 264908, 264510, 265008, 285007, 265008, 265009, 60432228, 33857402, 60433356, 265011, 87168559, 264600, 265017, 265018, 265019, 18108351, 264288, 284369, 21906766, 21906767, 21906788, 265020, 60170815, 264693, 65274620, 18108370, 264639, 18108384, 22278000, 264563, 18108390
3111	87754512 (6221, 6222)	Novel Protein sim. GBank gl3282231 (U75454) - C2H2 type zinc finger protein [Homo sapiens]	Contains protein domain (PF00008) - Zinc finger, C2H2 type	(transcript) factor	264488, 18108388, 66712502, 265017, 265018, 265019, 264448, 21906767, 265020, 33857023, 18108385, 18108388, 35698423, 52644332, 18108385, 18108388
3112	88043639 (6223, 6224)	Novel Protein sim. GBank gl3900848 (AC005023) - match to EST AA381117 (NID:2013436) [Homo sapiens]	Contains protein domain (PF00046) - Homeobox domain	homeobox	
3113	88207098 (6225, 6226)	Novel Protein sim. GBank gl2459910 (AF005856) - anon2A5 [Drosophila yakuba]		im7	18108397, 22278999, 264259, 29331824, 35896052, 264907, 264757, 60433438, 87168559, 264763, 264448, 18108354, 264288, 21908767, 21906769, 35695917, 264690, 264691, 264692, 264693, 18108365, 18108381, 18108384, 18108385, 18108388, 87168518, 22278000, 22279002
3114	79843167 (6227, 6228)	Novel Protein sim. GBank gl4886270[gb]AAB52281.2] - (U97002) similar to acyl-CoA dehydrogenases and epoxide hydrolases; Pfam domain PF00441 (Acyl-CoA_dh). Score=57.4, E-value=1.7e-16, N=2; contains similarity to Pfam domain PF00702 (Hydrolase). Score=57.4, E-value=1e-13, N=1 [C...]	Contains protein domain (PF00702) - haloacid dehalogenase-like hydrolase	hydrolase	
3115	94117986 (6229, 6230)	Novel Protein sim. GBank gl5032225[re]NP_005876.1pWBSC - Williams-Beuren syndrome chromosome region 11		transcript factor	60424179, 56182575, 264259, 29331824, 60424289, 29331828, 66712502, 264510, 265007, 60431735, 60433358, 55812038, 55811388, 285019, 264288, 264689, 21906769, 264691, 33857023, 264693, 60431528, 263974, 60431850, 56182323, 264559, 22279000, 22279002
3116	79842855 (6231, 6232)			UNCLASSIFIED	264905, 264758, 21906764, 264690
3117	87711288 (6233, 6234)			UNCLASSIFIED	264510, 265011, 18108351, 264288, 264689, 264691, 18108368, 18108372, 263981, 264558, 264564

3118	94685848 (6235, 6236)	Novel Protein sim. GBank gi 3880563 emb CAB01444.1 - (Z78018) predicted using Genefinder; similar to serine/threonine kinase; cDNA EST yk353d10.5 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00008) - Igf EGF-like domain	52645156, 52646842, 65274572, 56182575, 22278995, 56994075, 22278998, 35698286, 22278997, 22278998, 22278999, 284259, 29331822, 29331824, 68714117, 29331826, 29331827, 35696052, 29331828, 284905, 284908, 29331830, 52844045, 56182435, 284510, 284511, 265007, 265008, 265009, 284757, 52846317, 21906754, 33657084, 52844298, 87168474, 87168559, 265017, 265018, 264605, 265019, 264782, 284448, 284682, 264684, 264288, 264766, 56181582, 21906765, 21906766, 21906768, 21906769, 265020, 265022, 264690, 52844150, 264691, 33657023, 284693, 33657109, 33657349, 284628, 18108370, 60431528, 18108374, 35698423, 65274791, 60170394, 83373044, 87168518, 22279000, 22279002, 264486, 265006, 264288
3119	85728786 (6237, 6238)		Contains protein domain (PF00328) - Histidine acid phosphatase	264488, 264509, 264510, 264511, 264512, 264288, 264488
3120	87344040 (6239, 6240)	Novel Protein sim. GBank gi 5019819 gb AAD37863.1 AF14315 - (AF143152) putative NADH oxidoreductase complex I subunit [Caenorhabditis elegans]	UNCLASSIFIED	52644507, 52645156, 52846365, 52846842, 22278994, 56994075, 22278996, 22278999, 284259, 29331824, 29331827, 35698052, 52844045, 265008, 52846317, 87168474, 87168559, 21906765, 52844150, 33657023, 18108374, 264637
3121	94110735 (6241, 6242)	Novel Protein sim. GBank gi 4501877 ref NP_001088.1 pACR - acrosin	UNCLASSIFIED	18108392, 29331822, 29331824, 29331825, 284905, 265007, 55812038, 265019, 18108351, 264682, 264286, 264766, 21906764, 21906765, 21906768, 21906769, 55811957, 18108365, 18108366, 27486265, 18108374, 18108381, 18108384, 22279000, 22279002, 264482
3122	11814528 (6243, 6244)	Novel Protein sim. GBank gi 2439517 (AC002583) - putative RHO/RAC effector protein; 95% similarity to P49205 (PID.g1345860) [Homo sapiens]	Contains protein domain (PF00780) - CNH domain	264905
3123	88083003 (6245, 6246)		UNCLASSIFIED	56181686, 264259, 66714117, 60432288, 29331826, 29331827, 264907, 264908, 284828, 265009, 60433358, 33637402, 60433438, 264758, 18108351, 264286, 29148627, 29148629, 33657023, 33657109, 18108382, 56528488
3124	87768899 (6247, 6248)	Novel Protein sim. GBank gi 4980826 gb AAD35412.1 AE00171 - (AE001714) oxidoreductase, short chain dehydrogenase/reductase family [Thermotoga maritima]	Contains protein domain (PF00106) - short chain dehydrogenase	264905
3125	91216607 (6249, 6250)		UNCLASSIFIED	

3126	95337205 (6251, 6252)			UNCLASSIFIED	22278999, 264490, 264259, 60432049, 29331822, 60432289, 29146498, 52644045, 56182435, 265009, 60433438, 265010, 87188559, 265017, 265018, 55811150, 284763, 264683, 264369, 264685, 28148629, 33657023, 264693, 33657109, 18108374, 55811576, 18108385, 60432113, 22278902, 35696286, 22278996, 22278999, 26331826, 264908, 60433438, 87168559, 284604, 21908765, 21908769, 33657023, 33657349, 264629, 18108374, 18108377, 22279000, 22278902
3127	91639233 (6253, 6254)	Novel Protein sim. GBank gj 2828280 emb CAA16694.1 - (AL021887) putative protein [Arabidopsis thaliana]			
3128	87674330 (6255, 6256)	Novel Protein sim. GBank gj 3885828 (AF090133) - lin-7-A [Rattus norvegicus]	Contains protein domain (PF00395) - PDZ domain (Also known as DHR or GLGF).	misc_channel	22278998, 264259, 52644045, 265008, 21908754, 265017, 265018, 21908768, 18108376, 18108387, 22279000, 22279002
3129	87755412 (6257, 6258)	Novel Protein sim. GBank gj 3135273 (AC003058) - hypothetical protein [Arabidopsis thaliana]	Contains protein domain (PF00400) - WD domain, G-beta repeat	kinase	56182575, 264259, 29331825, 29331828, 52644045, 56182435, 60433356, 264600, 264682, 264763, 264764, 264369, 264288, 264688, 55811957, 264692, 33657023, 33657109, 60432113, 264564, 264566, 264638
3130	14983960 (6259, 6260)	Novel Protein sim. GBank gj 3329465 (AF064553) - NSD1 protein [Mus musculus]			
3131	95351469 (6261, 6262)	Novel Protein sim. GBank gj 1848277 (U86136) - telomerase-associated protein TP-1 [Homo sapiens]	Contains protein domain (PF00400) - WD domain, G-beta repeat	UNCLASSIFIED	56182575, 264259, 29331824, 264907, 56182435, 264594, 60433438, 55812038, 33109954, 21908754, 33657084, 87168474, 284448, 264768, 21908769, 55811957, 265020, 265021, 265022, 60170815, 33657023, 33657109, 33657182, 27486261, 33657349, 65274791, 60170394, 56182323, 83373044, 87168518, 264564

3132	95415459 (6263, 6264)	Novel Protein sim. GBank g j4680647 gb AAD27713.1 AF13293 - (AF132938) CGI-03 protein [Homo sapiens]	Contains protein domain (PF00789) - UBX domain	ubiquitin	52844507, 52646842, 52646365, 85274572, 56182575, 22278994, 22278995, 35698286, 5694075, 22278998, 22278997, 22278998, 22278999, 60432049, 52645080, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 29331828, 35698052, 284508, 52644045, 56182435, 284910, 60170831, 60432229, 60433356, 33857402, 55812038, 52646317, 21906754, 52644286, 85658542, 87168559, 285017, 265018, 285019, 284448, 264288, 264369, 52644229, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 60170815, 52644150, 264682, 33857023, 52645128, 33657109, 33657182, 27486261, 27486262, 27486265, 33657349, 35695763, 18108374, 18108376, 55811576, 35695855, 18108385, 18108387, 56526486, 87168518, 60432113, 22279002
3133	87379414 (6265, 6266)	Novel Protein sim. GBank g j4507613 ref NP_003738.1 pTNKS - TANKYRASE		polymerase	22278994, 22278998, 264905, 265006, 285007, 87168559, 264760, 21906767, 18108374, 22279000, 22279002, 264563 284595, 264369, 264685, 264628, 264568
3134	94649816 (6267, 6268)	Novel Protein sim. GBank g j1729827 sp P54633 ITALA_DICDI - FILOPODIN (TALIN HOMOLOG)			
3135	86389356 (6269, 6270)	Novel Protein sim. GBank g j3083478 (AF012927) - fibrinogen-binding protein [Streptococcus equi]		struct	22278998, 264095, 28331828, 33857402, 18108348, 263974
3136	94845839 (6271, 6272)	Novel Protein sim. GBank g j827101 p j S44092 - probable carrier protein c2 - Caenorhabditis elegans	Contains protein domain (PF00153) - Mitochondrial carrier proteins	transport	22278998, 264259, 264828, 265006, 265008, 60433438, 285018, 264764, 264288, 264768, 264689, 265020, 27486262, 263972, 85274791, 284557, 284558
3137	88257947 (6273, 6274)	Novel Protein sim. GBank g j3342730 (AC005331) - R31341_1 [Homo sapiens]		UNCLASSIFIED	22278995, 22278996, 22278997, 22278998, 22278999, 284259, 28331822, 29331825, 29331828, 28331827, 28331828, 284510, 285008, 21906754, 87168474, 265011, 87168559, 265017, 265018, 265019, 18108351, 284882, 264769, 21906765, 21906766, 21906767, 21906769, 55811957, 35695917, 265020, 265021, 52644150, 18108370, 18108374, 22279000, 22279002, 284482, 264488

3138	94130186 (6275, 6276)	Novel Protein sim. GBank gij4406759jgbjAAD20070j - (ACC00838) hypothetical protein [Arabidopsis thaliana]			284589, 284488, 284907, 284511, 284593, 33109854, 87188559, 284881, 284684, 284685, 284686, 284687, 284768, 284688, 284689, 284691, 284692, 284693, 33657109, 284631, 284634, 284635, 284636, 284637, 60170394, 83373044, 18108385, 18108388, 60432113, 22278000, 22278002, 22278997, 22278998, 22278999, 284905, 285018, 285019, 21906765, 285020, 284636, 284537
3139	87325503 (6277, 6278)	Novel Protein sim. GBank gij22893jprj1814452C - Hyp-rich glycoprotein [Zea mays]	UNCLASSIFIED		
3140	91222692 (6279, 6280)	Novel Protein sim. GBank gij832jembCAA37773j - (X53744) 68kDa subunit of signal recognition particle [Canis familiaris]	struct		22278995, 56994075, 35696286, 284908, 284909, 60433356, 21806754, 52644298, 87188474, 87188559, 284693, 284288, 284685, 284686, 285022, 284693, 27486282, 35695855, 284630, 284555, 284566
3141	87323564 (6281, 6282)	Novel Protein sim. GBank gij3213227 (AF035209) - putative v-SNARE VU1a [Mus musculus]	UNCLASSIFIED		56182575, 35696286, 28331828, 284909, 285009, 285018, 18108351, 284369, 21806766, 2848627, 285020, 284628, 284629, 284631, 18108385
3142	85418028 (6283, 6284)	Novel Protein sim. GBank gij2488197jsgj95245[C561_PIG - CYTOCHROME B561 (CYTOCHROME B-561)]	cytochrome		52645156, 52646365, 22278995, 35696286, 22278998, 22278999, 60432049, 284259, 28331822, 28331824, 28331827, 28148498, 56182435, 285007, 80170831, 60432229, 33657402, 284595, 60433438, 284758, 21906754, 284288, 284768, 284687, 32844228, 21908765, 21908767, 21908768, 60170815, 52644150, 65274620, 33657109, 35695763, 18108370, 18108376, 65274791, 35695855, 284631, 284557, 87168518, 60432113, 22278000
3143	95351475 (6285, 6286)	Novel Protein sim. GBank gij5420387jembjCAB46878.1j - (AJ243459) proteophosphoglycan [Leishmania major]	UNCLASSIFIED		284488, 56182575, 22278996, 22278998, 22278999, 28331822, 28331824, 60432289, 35696052, 28331828, 284508, 284905, 284906, 284907, 284908, 284909, 52644045, 56182435, 284511, 284512, 285008, 284910, 60432229, 33657402, 60433356, 60433438, 55812038, 285011, 285019, 284760, 284763, 284448, 284764, 284884, 284288, 284685, 284686, 284768, 284889, 21906765, 21908768, 21908767, 35695917, 284690, 33657023, 284693, 283987, 33657109, 284628, 284629, 18108374, 283976, 55811576, 35695855, 284630, 284631, 284632, 284634, 284635, 284636, 284637, 284558, 87168518, 60432113, 22278000, 22278002, 284563, 284566, 284488

3144	95336329 (6287, 6288)	Novel Protein sim. GBank gi4884468[embjCAB43322.1] - (AL050225) hypothetical protein [Homo sapiens]			264488, 18108398, 22278998, 35698286, 22278997, 29331826, 29331827, 35696052, 29331828, 264106, 265008, 265007, 265009, 33857402, 85658542, 265011, 18108351, 284448, 264369, 21908765, 21908766, 21908767, 265020, 265021, 52844150, 27486261, 18108370, 18108374, 35698423, 58182323, 83373044, 22279000, 22279002, 264567
3145	86611657 (6289, 6290)	Novel Protein sim. GBank gi3878709[embjCAB03330] - (Z81118) Similarity to Human endosomal protein P182 (TR:Q15075); cDNA EST EMBL:Z14487 comes from this gene; cDNA EST EMBL:Z14556 comes from this gene; cDNA EST EMBL:D27011 comes from this gene; cDNA EST EMBL:D27015 comes from t...	UNCLASSIFIED		18108397, 29331824, 29146499, 20281100, 265006, 55812038, 265010, 21908766, 29148627, 21908769, 29148784, 264892, 33657023, 33657109, 35695763, 263981, 58182323, 87168518
3146	87756314 (6281, 6292)	Novel Protein sim. GBank gi2135746[pirjS69890] - mitogen inducible gene mlg-2 - human	Contains protein domain (PF00169) - PH domain	- struct	264259, 29331826, 29331828, 29331830, 264510, 264511, 265007, 265009, 264600, 265017, 18108351, 284448, 264369, 21908768, 265021, 264892, 33657109, 18108374, 35698423, 35695855, 60432113, 264564
3147	94848512 (6283, 6294)	Novel Protein sim. GBank gi3874278[embjCAB07315.1] - (Z92825) predicted using GeneFinder; cDNA EST yk315e12.3 comes from this gene; cDNA EST yk315e12.5 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00702) - haloacid dehalogenase-like hydrolase	UNCLASSIFIED	58181866, 35698286, 60432049, 264259, 58182181, 29331825, 60432289, 35696052, 58182435, 265008, 284910, 60431735, 60433356, 60433438, 265010, 284448, 264288, 265022, 33657023, 33657109, 60431528, 65274791, 264631, 56182323, 264404, 22279002
3148	95362169 (6295, 6296)	Novel Protein sim. GBank gi5225322[gbjAAD40851.1]AF08310 - (AF083108) sirutin type 3 [Homo sapiens]		UNCLASSIFIED	35698286, 35698052, 264511, 85658542, 87168474, 264764, 35696423, 264555, 264556, 264557, 264558, 83373044, 58526488, 60432113
3149	95308548 (6297, 6298)	Novel Protein sim. GBank gi4200446 (AF102777) - FYVE finger-containing phosphoinositide kinase [Mus musculus]	Contains protein domain (PF01363) - FYVE zinc finger	- eph	29331822, 35698052, 264109, 29148629, 18108381
3150	87655472 (6299, 6300)	Novel Protein sim. GBank gi3378454[embjCAA76893] - (Y17850) ganglioside-induced differentiation associated protein 1 [Mus musculus]	Contains protein domain (PF00043) - Glutathione S-transferases	- transferase	264259, 29331822, 29331824, 29331825, 29331827, 52846317, 264686, 35695855, 58182323, 264639
3151	87772355 (6301, 6302)	Novel Protein sim. GBank gi172591 (M63577) - SFP1 [Saccharomyces cerevisiae]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	- oncogene	29331822, 265008
3152	85698108 (6303, 6304)			UNCLASSIFIED	21908754, 87168559, 264605, 21908768, 52844150, 27486264, 35698423, 22279000

3153	95317298 (6305, 6306)	Novel Protein sim. GBank gi 4895041 gb AAD32705.1 AF14395 - (AF143957) coronin-3 [Mus musculus]	Contains protein domain (PF00400) - struct WD domain, G-beta repeat		264488, 5264365, 35688286, 22278996, 22278997, 22278999, 60432049, 264259, 29331826, 60432289, 33656970, 264508, 264908, 33657402, 264595, 60433438, 87188474, 87188559, 264601, 265019, 264448, 264682, 264764, 264288, 264369, 264766, 21908765, 21906766, 21906767, 21908768, 21908769, 29148784, 265021, 265022, 60170815, 52644150, 264690, 264691, 33657023, 65274620, 33657109, 18108370, 35695855, 264636, 60170394, 87188518, 60432113, 22278900, 22279002, 22278998, 264259, 29331824, 66712502, 265008, 265010, 265017, 18108354, 264691, 33657023, 264693, 20281149, 18108374, 29331828, 264509, 264905, 264908, 264510, 264511, 264512, 33657402, 264681, 264683, 33657023, 18108370, 264634, 264639, 18108385, 264563, 264486
3154	87718573 (6307, 6308)	Novel Protein sim. GBank gi 4880681 gb AAD27720.1 AF13294 - (AF132945) CGI-11 protein [Homo sapiens]	ATPase associated		59182575, 22278996, 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 66714117, 29331825, 29331826, 29331827, 35696052, 52644045, 265007, 265008, 60170831, 60432229, 60433356, 21908754, 33109954, 87168474, 265010, 265017, 265018, 265019, 18108351, 264448, 264288, 264689, 21906766, 21908768, 21908769, 35695917, 265020, 265022, 264692, 18108370, 35696423, 56182323, 22279002
3155	87762384 (6309, 6310)	Novel Protein sim. GBank gi 728837 sp P39194 ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII	UNCLASSIFIED		18108396, 264259, 29331826, 35696052, 29148498, 87188559, 265017, 264448, 264288, 264691, 18108368, 52645129, 35696423, 52644332
3156	87737449 (6311, 6312)	Novel Protein sim. GBank gi 5630078 gb AAD45821.1 AC00601 - (AC006017) N-acetylgalectosaminyltransferase; similar to Q10473 (PID:g1708559) [Homo sapiens]	Contains protein domain (PF00652) - Similarity to lectin domain of ricin beta-chain, 3 copies.	transferase	264488, 263974
3157	88259577 (6313, 6314)				
3158	80034118 (6315, 6316)	Novel Protein sim. GBank gi 5306064 gb AAD41895.1 AF15677 - (AF156778) ASB-3 protein [Homo sapiens]	Contains protein domain (PF00023) - Ank repeat	kinase	
3159	84124114 (5317, 6318)	Novel Protein sim. GBank gi 5531272 emb CAB50897.1 - (AJ243800) WSC4 homologue [Kluyveromyces fragilis]	UNCLASSIFIED		56182575, 22278999, 29331824, 264106, 60433356, 264758, 265011, 87188559, 264448, 18108354, 264768, 21908768, 265020, 264691, 264692, 33657109, 18108374, 35696423, 264555, 60170394, 22279000
3160	80221068 (6319, 6320)	Novel Protein sim. GBank gi 3930525 (AF084447) - sex-determination protein homolog Fem1a [Mus musculus]	Contains protein domain (PF00023) - Ank repeat	struct	18108351, 264555, 264556, 264557, 264558, 264559

3161	88074111 (6321, 6322)				264488, 2227895, 2227897, 2227898, 264259, 29331822, 60432289, 29331828, 52844045, 265017, 265018, 264448, 264288, 21908784, 21908767, 265020, 18108374, 264638, 264588
------	-----------------------	--	--	--	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Table 2

Tissue ID	Tissue Name	Tissue Information	Disease Association
20281069	192xN	Protein-protein Interactions	Any
20281071	192xN	Protein-protein Interactions	Any
20281149	192xN	Protein-protein Interactions	Any
20281152	192xN	Protein-protein Interactions	Any
264111	276xN	Protein-protein Interactions	Any
264112	276xN	Protein-protein Interactions	Any
263966	384xN	Protein-protein Interactions	Any
263967	384xN	Protein-protein Interactions	Any
264110	552xN	Protein-protein Interactions	Any
18108379	SPH 52.1 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,
18108381	SPH 52.2 (Fetal Lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
18108383	SPH 52.3 (B's Lymphoma- Raji)	Lymphoma derived from B cells	Blood cancers, hematopoiesis, leukemia
18108368	SPH 52.5 (Salivary Gland)	Salivary Gland	Dry mouth, infection
18108384	SPH 52.6 (Brain- Thalamus)	Thalamus	Brain cancer, head injury, obesity, neurological disorders, neuropsychiatric disorders
18108394	SPH 53.1 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,
18108355	SPH 53.2 (Fetal Lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
18108359	SPH 53.3 (B's Lymphoma- Raji)	Lymphoma derived from B cells	Blood cancers, hematopoiesis, leukemia
18108361	SPH 53.4 (Mammary Gland)	Mammary Gland	Lactation disorders, breast cancer
18108362	SPH 53.5 (Salivary Gland)	Salivary Gland	Dry mouth, infection
18108366	SPH 53.6 (Brain- Thalamus)	Thalamus	Brain cancer, head injury, obesity, neurological disorders, neuropsychiatric disorders
18108354	SPH 54.1 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,
18108392	SPH 54.2 (Fetal Lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
18108348	SPH 54.3 (B's Lymphoma- Raji)	Lymphoma derived from B cells	Blood cancers, hematopoiesis, leukemia
18108382	SPH 54.4 (Mammary Gland)	Mammary Gland	Lactation disorders, breast cancer
18108395	SPH 54.5 (Salivary Gland)	Salivary Gland	Dry mouth, infection
18108365	SPH 54.6 (Brain- Thalamus)	Thalamus	Brain cancer, head injury, obesity, neurological disorders, neuropsychiatric disorders
18108397	SPH 55.1 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,
18108398	SPH 55.2 (Fetal Lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
18108364	SPH 55.3 (B's Lymphoma- Raji)	Lymphoma derived from B cells	
18108388	SPH 55.4 (Mammary Gland)	Mammary Gland	Lactation disorders, breast cancer
18108358	SPH 55.5 (Salivary Gland)	Salivary Gland	Dry mouth, infection
20281099	SPH 56.2 (MG63)		
20281100	SPH 56.3 (U1SMC)		
264404	SPH.1 (Brain)	Whole Brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection

264510	SPH.10 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264511	SPH.11 (Placenta)	Placenta	Infertility, birth defects
264512	SPH.12 (Thyroid)	Thyroid	Hyperparathyroidism, Hypoparathyroidism
264555	SPH.13 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264556	SPH.14 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264557	SPH.15 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264558	SPH.16 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264559	SPH.17 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264569	SPH.19 (One Fetal tissue and two cell lines)	Mixed	
264687	SPH.19.1 (fetal thymus - CRL7046)	Fetal Thymus	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, immunodeficiencies
264688	SPH.19.2 (hematopoietic stem cells - CRL2043)	Hematopoietic stem cells	Leukemia, osteoporosis, post-chemotherapeutic stem cell repopulation
264689	SPH.19.3 (osteogenic sarcoma cell lines - HTB96)	Osteogenic Sarcoma	Sarcomas, osteoporosis, osteopetrosis
264690	SPH.19.4 (Fetal Liver)	Fetal liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
264691	SPH.19.5 (Heart)	Heart	Cardiomyopathy, Atherosclerosis, Hypertension, Congenital heart defects, Aortic stenosis, Atrial septal defect (ASD), Atrioventricular (A-V) canal defect, Ductus arteriosus, Pulmonary stenosis, Subaortic stenosis, Ventricular septal defect (VSD), valve diseases, Tuberous sclerosis, Scleroderma, Obesity, Transplantation
264692	SPH.19.6 (Spleen)	Spleen	Hemophilia, Hypercoagulation, Idiopathic thrombocytopenic purpura, Immunodeficiencies, Graft versus host
264693	SPH.19.7 (Pituitary)	Pituitary	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264482	SPH.2 (Brain)	Brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection

264600	SPH.21 (Fetal Brain)	Fetal brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264601	SPH.22 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264602	SPH.23 (Thyroid)	Thyroid	Hyperthyroidism and Hypothyroidism
264603	SPH.24 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264604	SPH.25 (Lymph Node)	Lymph Node	Lymphedema, Allergies
264605	SPH.26 (Placenta)	Placenta	Infertility, birth defects
264634	SPH.28 (Heart)	Heart	Cardiomyopathy, Atherosclerosis, Hypertension, Congenital heart defects, Aortic stenosis, Atrial septal defect (ASD), Atrioventricular (A-V) canal defect, Ductus arteriosus, Pulmonary stenosis, Subaortic stenosis, Ventricular septal defect (VSD), valve diseases, Tuberous sclerosis, Scleroderma, Obesity, Transplantation
264635	SPH.29 (Fetal Kidney)	Fetal Kidney	Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome
264483	SPH.3 (Bone Marrow)	Bone marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264636	SPH.30 (Lymph Node)	Lymph Node	Lymphedema, Allergies
264637	SPH.31 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264638	SPH.32 (Thyroid)	Thyroid	Hyperthyroidism and Hypothyroidism
264639	SPH.33 (Fetal Brain)	Fetal brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264484	SPH.4 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264758	SPH.44.1 (Kidney)	Kidney	Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome
264760	SPH.44.2 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation

264762	SPH.44.3 (Heart)	Heart	Cardiomyopathy, Atherosclerosis, Hypertension, Congenital heart defects, Aortic stenosis, Atrial septal defect (ASD), Atrioventricular (A-V) canal defect, Ductus arteriosus, Pulmonary stenosis, Subaortic stenosis, Ventricular septal defect (VSD), valve diseases, Tuberous sclerosis, Scleroderma, Obesity, Transplantation
264764	SPH.44.4 (Prostate)	Prostate	Prostate Cancer
264766	SPH.44.5 (Spleen)	Spleen	Hemophilia, Hypercoagulation, Idiopathic thrombocytopenic purpura, Immunodeficiencies, Graft versus host
264768	SPH.44.6 (pituitary)	Pituitary	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264769	SPH.44.7 (Uterus)	Uterus	Infertility, birth defects
264905	SPH.48.1 (Burkitt's Lymphoma- Raji)	Burkitt's Lymphoma	Lymphoma, blood cancers
264906	SPH.48.2 (Thalamus- Brain)	Thalamus	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264907	SPH.48.3 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,
264908	SPH.48.4 (Fetal Lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
264909	SPH.48.5 (Salivary Gland)	Salivary Gland	Dry mouth, infection
264910	SPH.48.6 (Mammary Gland)	Mammary Gland	Lactation disorders, breast cancer
265006	SPH.50.1 (B's lymphoma)	Burkitt's Lymphoma	Lymphoma, blood cancers
265007	SPH.50.2 (thalamus)	Thalamus	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
265008	SPH.50.3 (adrenal gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,
265009	SPH.50.4 (fetal lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
265010	SPH.50.5 (salivary gland)	Salivary Gland	Dry mouth, infection
265011	SPH.50.6 (mammary gland)	Mammary Gland	Lactation disorders, breast cancer
18108385	SPH.51.1 (MCF-7)	Breast Cancer	Breast Cancer
18108370	SPH.51.2 (CCRF-CEM)	Cancer Cell line	Cancer
18108374	SPH.51.3 (K-562)	Cancer Cell line	Cancer
18108351	SPH.51.4 (OVCAR-3)	Ovarian cancer	Ovarian cancer
18108372	SPH.51.5 (HL-60)	Cancer Cell line	Cancer
264486	SPH.6 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host.

264508	SPH.8 (Fetal Brain)		Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264509	SPH.9 (Lymph Node)	Lymph Node	Lymphedema , Allergies
20798451	SRH.56.3 (U1SMC)		
264487	SRH.1 (Brain)	Brain	Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264534	SRH.11 (Bone marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264535	SRH.12 (Bone marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264563	SRH.19 (Fetal Brain)	Fetal brain	Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264488	SRH.2 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264564	SRH.20 (Lymph Node)	Lymph Node	Lymphedema , Allergies
264565	SRH.21 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264566	SRH.22 (Placenta)	Placenta	Infertility, birth defects
264567	SRH.23 (Thyroid)	Thyroid	Hyperthyroidism and Hypothyroidism
264591	SRH.25 (Fetal Brain)	Fetal brain	Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264592	SRH.26 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264593	SRH.27 (thyroid)	Thyroid	Hyperthyroidism and Hypothyroidism
264594	SRH.28 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264595	SRH.29 (Lymph Node)	Lymph Node	Lymphedema , Allergies
264489	SRH.3 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,

264596	SRH.30 (Placenta)	Placenta	Infertility, birth defects
264628	SRH.33 (Fetal Kidney)	Fetal kidney	Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome
264629	SRH.34 (Lymph Node)	Lymph Node	Lymphedema, Allergies
264630	SRH.35 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264631	SRH.36 (Thyroid)	Thyroid	Hyperthyroidism and Hypothyroidism
264632	SRH.37 (Fetal Brain)	Fetal Brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264490	SRH.4 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host
264681	SRH.43.1 (fetal thymus - CRL7046)	Fetal Thymus	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, immunodeficiencies
264682	SRH.43.2 (hematopoietic stem cells - CRL2043)	Hematopoietic stem cells	Leukemia, osteoporosis, post-chemotherapeutic stem cell repopulation
264683	SRH.43.3 (osteogenic sarcoma cell lines - HTB96)	Osteogenic Sarcoma	Sarcomas, osteoporosis, osteopetrosis
264684	SRH.43.4 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
264685	SRH.43.6 (Spleen)	Spleen	Hemophilia, Hypercoagulation, Idiopathic thrombocytopenic purpura, Immunodeficiencies, Graft versus host
264686	SRH.43.7 (pituitary)	Pituitary	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264757	SRH.44.1 (Kidney)	Kidney	Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome
264759	SRH.44.2 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
264761	SRH.44.3 (Heart)	Heart	Cardiomyopathy, Atherosclerosis, Hypertension, Congenital heart defects, Aortic stenosis, Atrial septal defect (ASD), Atrioventricular (A-V) canal defect, Ductus arteriosus, Pulmonary stenosis, Subaortic stenosis, Ventricular septal defect (VSD), valve diseases, Tuberous sclerosis, Scleroderma, Obesity, Transplantation
264763	SRH.44.4 (Prostate)	Prostate	Prostate Cancer
264765	SRH.44.5 (Spleen)	Spleen	Hemophilia, Hypercoagulation, Idiopathic thrombocytopenic purpura, Immunodeficiencies, Graft versus host

264767	SRH.44.6 (Pituitary)	Pituitary	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264828	SRH.46.1 (Lymph Node)	Lymph Node	Lymphedema, Allergies
264887	SRH.47.5 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
18108377	SRH.50.1 (B's lymphoma)	Burkitt's Lymphoma	Lymphoma, blood cancers
18108380	SRH.50.2 (thalamus)	Thalamus	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
18108396	SRH.50.3 (adrenal gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,
18108391	SRH.50.4 (fetal lung)	Fetal Lung	Airway diseases, infection
18108357	SRH.50.5 (salivary gland)	Salivary Gland	Dry mouth, infection
18108390	SRH.50.6 (mammary gland)	Mammary Gland	Lactation disorders, breast cancer
264532	SRH.9 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
263974	736xN		
263976	736xN		
263981	736xN		
20281166	96xN		
20281169	96xN		
20281171	96xN		
263994	cDNA-ORF Selection		
264080	Mx96		
21906754	NQH 6.1 (HH729)		
22278996	NQH 6.10 (PrEC)	Endothelial cells	heart disease, cancer
22278997	NQH 6.11 (CAEC)	Endothelial cells	heart disease, cancer
22278998	NQH 6.12 (CSC)	Cancer Cell line	Cancer
22278999	NQH 6.13 (NHNPC)	Cancer Cell line	Cancer
22279000	NQH 6.14 (NHMC-RM)	Cancer Cell line	Cancer
22279002	NQH 6.15 (Hypothalamus)	Hypothalamus	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection, Obesity
21906764	NQH 6.2 (In Dated Platelets)	Platelets	Clotting diseases, stroke
21906765	NQH 6.3 (HuVec)	Endothelial cells	heart disease, cancer
87168474	NQH 6.3 (Sized-HUVEC)	Endothelial cells	heart disease, cancer
21906766	NQH 6.4 (U8MVEC-myo)	Cancer Cell line	Cancer
21906767	NQH 6.5 (NHMC-neo)	Cancer Cell line	Cancer
21906768	NQH 6.6 (NHEK)	Cancer Cell line	Cancer
21906769	NQH 6.7 (ByCAEC)	Endothelial cells	heart disease, cancer
22278994	NQH 6.8 (NHA)	Cancer Cell line	Cancer

22278995	NQH 6.9 (PrSC)	Cancer Cell line	Cancer
27486261	NQH 7.1 (Jurkat E6-untreated)	Cancer Cell line	Cancer
27486262	NQH 7.2 (TF1-untreated)	Cancer Cell line	Cancer
27486264	NQH 7.3 (U87-untreated)	Cancer Cell line	Cancer
27486265	NQH 7.4 (THP1-untreated)	Cancer Cell line	Cancer
29331822	NQH 8.1 (Brain- amygdala)		Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
29331824	NQH 8.2 (Brain-hippocampus)		Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
29331825	NQH 8.3 (Brain- substantia nigra)		Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
29331826	NQH 8.4 (small intestine)	Small intestine	digestive diseases, obesity, diabetes
29331827	NQH 8.5 (Spinal cord)	Spinal chord	paralysis, neurodegenerative disorders
29331828	NQH 8.6 (stomach)	Stomach	Stomach cancer
29331830	NQH 8.7 (Trachea)	Trachea	Airway diseases, infection
87168518	NQH 9.1 (Sized-MG-63_treatment pool)		
87168559	NQH 9.2 (Sized-HEPG2 untreated)		
35695763	NQH.10.1 (MCF-7untreated)	Cancer Cell line	Cancer
35695855	NQH.10.2 (U-937_treatment pool)	Cancer Cell line	Cancer
35695917	NQH.10.3 (JAR)	Cancer Cell line	Cancer
35696052	NQH.10.4 (PA-1)	Cancer Cell line	Cancer
35696286	NQH.10.5 (CADMEC)	Endothelial cells	heart disease, cancer
35696423	NQH.10.6 (CADMEC_LA)	Endothelial cells	heart disease, cancer
52644045	NQH.11.1 (SK-PN-DW)	Cancer Cell line	Cancer
52644150	NQH.11.2 (Chorionic Villus Cells)	Chorionic villus	fertility, birth defects
52644229	NQH.11.3 (A549)	Cancer Cell line	Cancer
52644296	NQH.11.4 (U266B1)	Cancer Cell line	Cancer
52644332	NQH.11.5 (Daoy)	Cancer Cell line	Cancer
52644507	NQH.11.6 (SW1783)	Cancer Cell line	Cancer
52645080	NQH.12.1 (U-118MG)	Cancer Cell line	Cancer
52645129	NQH.12.2 (A204)	Cancer Cell line	Cancer
52645156	NQH.12.3 (T24)	Cancer Cell line	Cancer
52646317	NQH.12.4 (G-401)	Cancer Cell line	Cancer
52646365	NQH.12.5 (CaSki)	Cancer Cell line	Cancer
52646842	NQH.12.6 (SHP-77)	Cancer Cell line	Cancer

60424179	NQH.14.1 (Yale75_breast carcinoma)	Breast carcinoma	Breast Cancer
60424269	NQH.14.2 (Yale78B_ovarytumor)	Ovary tumor	Ovarian cancer
60431528	NQH.14.3 (Yale79_prostateBPH)	Prostate	Prostate Cancer
60431602	NQH.14.4 (Yale80_ProstateAdenocarcinoma)	Prostate	Prostate Cancer
60431735	NQH.14.5 (Yale86_UterineMyoma)	Uterine Myoma	Uterine Cancer
60431850	NQH.14.6 (Yale207_Myometrium)	Myometrium	Fertility
60432049	NQH.15.1 (Yale99_cervix)	Cervix	Osteoporosis, cervical cancer
60432113	NQH.15.2 (Yale45_spleenITP)		Hemophilia, Hypercoagulation, Idiopathic thrombocytopenic purpura, Immunodeficiencies, Graft versus host
60432229	NQH.15.3 (Yale16_Skin)	Skin	wound healing, melanoma
60432289	NQH.15.4 (Yale137_Parotid)		
60433356	NQH.15.5 (Yale38_SmallIntestine)	Small intestine	digestive diseases, obesity, diabetes
60433438	NQH.15.6 (Yale28_ColonAscending)	Colon	Colon cancer
65274444	NQH.17.1 (Larynx)	Larynx	Cancer
65274572	NQH.17.2 (Duodenum)	Duodenum	
65274620	NQH.17.3 (Kidney, Primary tumors)		Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome
65274727	NQH.17.4 (Lung Pleura, normal)	Lung	Airway diseases, infection
65274791	NQH.17.5 (Lung, Normal Adult)	Lung	Airway diseases, infection
83373044	NQH.18.230 (Pooled adrenal gland, placenta)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,
85658542	NQH.18.560 (Pooled uterus, BeWo pool)	Uterus	Infertility, birth defects
33656970	NQH.9.1 (MG-63_treatment pool)	Cancer Cell line	Cancer
33657023	NQH.9.2 (HEPG2 untreated)		Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
33657084	NQH.9.3 (PC3_untreated)	Cancer Cell line	Cancer
33657109	NQH.9.4 (TF-1_TPA)	Cancer Cell line	Cancer
33657182	NQH.9.5 (TF-1_TPO)	Cancer Cell line	Cancer
33657349	NQH.9.6 (TF-1_Hemin)	Cancer Cell line	Cancer
33657402	NQH.9.7 (HFDPC)	Cancer Cell line	Cancer
264259	NQH1 (Mixture of eight adult & two fetal tissues)		
264288	NQH2 (Ten tissues plus lymphocyte control)		
264448	NQH3 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
265017	NQH4.1 (lymph node)	Lymph Node	Lymphedema, Allergies

265018	NQH4.2 (fetal kidney)	Fetal Kidney	Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome
66712502	NQH4.2 (Sized)		
265019	NQH4.3 (pituitary gland)		Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection, Obesity
66714117	NQH4.3 (Sized)		
265020	NQH4.4 (testis)	testis	Infertility, birth defects
265021	NQH4.5 (fetal liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
265022	NQH4.6 (thyroid)	Thyroid	Hyperthyroidism and Hypothyroidism
18108376	NQH5.1 (MCF-7)	Breast cancer	Breast Cancer
18108387	NQH5.2 (CCRF-CEM)	Cancer Cell line	Cancer
264952	NRL1: HPLC FRACTIONATION OF RE-LIG		
263971	Old BB3 Baits		
263969	Old BB5 Baits		
263975	ORFSEL		
263972	OTHER Baits		
263978	pGALORF		
264106	PPBAITS		
264088	QC-YA7		
264089	QC-YA8		
264102	Resequenced Interactors		
264369	RRH.1		
60170394	RRH.10.1 (MCF-7untreated)	Breast cancer	Breast Cancer
60170615	RRH.10.2 (U-937_treatment pool)	Cancer Cell line	Cancer
60170831	RRH.10.3 (JAR)	Cancer Cell line	Cancer
60174639	RRH.11.8 (HeLa)	Cancer Cell line	Cancer
264113	rrQEA Baits		
263973	RRQEA_B5 baits		
29146498	SRD 3.1 (SKMC)	Cancer Cell line	Cancer
29146499	SRD 3.2 (SKMC)	Cancer Cell line	Cancer
29147620	SRD 3.3 (RPTEC)	Cancer Cell line	Cancer
29148627	SRD 3.4 (HRCE)	Cancer Cell line	Cancer
29148629	SRD 3.6 (HRE)	Cancer Cell line	Cancer
29148784	SRD 3.7 (HRE)	Cancer Cell line	Cancer
55810764	SRD.7.1 (Lymph Node)	Lymph Node	Lymphedema, Allergies
55811150	SRD.7.2 (pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
55811386	SRD.7.3 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,

55811576	SRD.7.4 (Pituitary Gland)	Pituitary	Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection, Obesity
55811957	SRD.7.5 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
55812038	SRD.7.6 (Fetal Kidney)	Fetal kidney	Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome
56181562	SRD.8.1 (Lymph Node)	Lymph Node	Lymphedema , Allergies
56181686	SRD.8.2 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
56182181	SRD.8.3 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy , Congenital Adrenal Hyperplasia,
56182323	SRD.8.4 (Pituitary Gland)	Pituitary	Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection, Obesity
56182435	SRD.8.5 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
56182575	SRD.8.6 (Fetal Kidney)		Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome
32833986	SRD4: HL adapter		
56526486	SRD5.1:rr fragments		
33109954	SRD5: long-RXRJ		
56994075	SRD9.1 (CS/SC)	Cancer Cell line	Cancer
263977	TSC Screen 1		

Disclosed Sequences of ORFX Nucleic Acid and Polypeptide Sequences

<210> 1
 <211> 312
 <212> DNA
 <213> Homo sapiens

<400> 1
 gtcgacagtg tcatgggagc cgcaccggcg atcggcggtg cgagcacgtt cggtagccgc
 60
 agaggaatcg gtaccgcgcg acgaggagga cgggcgggca gtcggcgac gtaggtgccg
 120
 gcgagcnncg atccgtctgc ctcccacca ccatctcgg cagccaggtt gtcgaggatg
 180
 tcgctgggtg agcgtctgc cgcgtctgct tcttgctcag ccgtcttgcg ggactgggac
 240
 tgctcctcgg ccttggtcag ggcagcgtca atgtcggcga aatccgattc tgccttcggc
 300
 tcagagacgc gt
 312

<210> 2
 <211> 104
 <212> PRT
 <213> Homo sapiens

<400> 2
 Val Asp Ser Val Met Gly Ala Arg Pro Ala Ile Gly Val Ala Ser Thr
 1 5 10 15
 Phe Gly Thr Ala Arg Gly Ile Gly Thr Ala Arg Arg Gly Gly Arg Ala
 20 25 30
 Gly Ser Ser Ala Thr Ser Gly Pro Ala Ser Xaa Asp Pro Ser Ala Ser
 35 40 45
 Pro Pro Pro Ile Ser Ala Ala Arg Leu Ser Arg Met Ser Leu Val Glu
 50 55 60
 Arg Cys Ser Ala Ser Ala Ser Cys Ser Ala Val Leu Arg Asp Trp Ala
 65 70 75 80
 Cys Ser Ser Ala Leu Val Arg Ala Ala Ser Met Ser Ala Lys Ser Asp
 85 90 95
 Ser Ala Phe Gly Ser Glu Thr Arg
 100

<210> 3
 <211> 987
 <212> DNA
 <213> Homo sapiens

<400> 3
 gagctcggta cccccgctcg cgccgtcaag gcgcaggtca aagtgggtgg ccgtggcaag
 60
 gctggcgggtg taaagatcgc taagtccccg gccgaggcgg ctgaacgcgc cgaggccatc
 120
 ctccggtatgg acatcaaggg tcacaccggt cacaaggtga tggtcgctga aggtgctgac
 180
 attgccgagg aatactactt ctcgatcttg ctcgaccgtg gagagcgtcg ctaccttgcg
 240
 atgtgctcgc gtgaggggtg catggacatc gagacccttg ctaaggagcg ccccgagggt
 300
 ctccccaagg tgccggtcga cccgattgac ggcgttgacg atgctaaagc ccgcgaaatc
 360
 cttagcgagg caggggtccc cgactctgag caggacgcta tcgtcccggc tgttctcaag
 420
 ctgtgggaga cctaccgtga cgaggatgcc accctcgctg aggtcaaccc gatgatcaag
 480
 accggcgatg gacgcaccc ctggatcgac ggcaagatga ctgttgacaa caacgcatcc
 540
 ttccgccagc ctgaccgcgc cggcttggtg gatcgcgcca ccaccgaccc gctcgagttg
 600
 cgtgccggcg agctcgggtc caactacgtc aagcttgacg gcaacgtcgg cgtcatcgga
 660
 aacgggtcgt gtctggteat gacgaccctt gactgcgttg cgtacgccgg tgagaacttc
 720
 ccgggatctc cagctccgcg taacttcctc gacatcggtg gcggcgctc ggccgagatc
 780
 atggccaacg gccttgacct catcatgagt gacgagcagg tgcgttcctg gtccgtcaac
 840
 gtctttggcg gtatcaccgc ctgtgaccag gtggcgcttg gtatcaaggg cgctctcgag
 900
 aagttgggcy acaaggccgt gaagccgctc gtcgtccgtc tggacggcaa cgctgtggcc
 960
 gagggcagaa agattctcga ggaattc
 987

<210> 4
 <211> 329
 <212> PRT
 <213> Homo sapiens

<400> 4
 Glu Leu Gly Thr Pro Val Val Ala Val Lys Ala Gln Val Lys Val Gly
 1 5 10 15
 Gly Arg Gly Lys Ala Gly Gly Val Lys Ile Ala Lys Ser Pro Ala Glu
 20 25 30
 Ala Ala Glu Arg Ala Glu Ala Ile Leu Gly Met Asp Ile Lys Gly His
 35 40 45

Thr Val His Lys Val Met Val Ala Glu Gly Ala Asp Ile Ala Glu Glu
 50 55 60
 Tyr Tyr Phe Ser Ile Leu Leu Asp Arg Gly Glu Arg Arg Tyr Leu Ala
 65 70 75 80
 Met Cys Ser Arg Glu Gly Gly Met Asp Ile Glu Thr Leu Ala Lys Glu
 85 90 95
 Arg Pro Glu Ala Leu Ala Lys Val Pro Val Asp Pro Ile Asp Gly Val
 100 105 110
 Asp Asp Ala Lys Ala Arg Glu Ile Leu Ser Glu Ala Gly Phe Pro Asp
 115 120 125
 Ser Glu Gln Asp Ala Ile Val Pro Ala Val Leu Lys Leu Trp Glu Thr
 130 135 140
 Tyr Arg Asp Glu Asp Ala Thr Leu Val Glu Val Asn Pro Met Ile Lys
 145 150 155 160
 Thr Gly Asp Gly Arg Ile Leu Ala Ile Asp Gly Lys Met Thr Val Asp
 165 170 175
 Asn Asn Ala Ser Phe Arg Gln Pro Asp Arg Ala Gly Leu Val Asp Arg
 180 185 190
 Ala Thr Thr Asp Pro Leu Glu Leu Arg Ala Gly Glu Leu Gly Leu Asn
 195 200 205
 Tyr Val Lys Leu Asp Gly Asn Val Gly Val Ile Gly Asn Gly Ala Gly
 210 215 220
 Leu Val Met Ser Thr Leu Asp Cys Val Ala Tyr Ala Gly Glu Asn Phe
 225 230 235 240
 Pro Gly Ser Pro Ala Pro Ala Asn Phe Leu Asp Ile Gly Gly Gly Ala
 245 250 255
 Ser Ala Glu Ile Met Ala Asn Gly Leu Asp Leu Ile Met Ser Asp Glu
 260 265 270
 Gln Val Arg Ser Val Phe Val Asn Val Phe Gly Gly Ile Thr Ala Cys
 275 280 285
 Asp Gln Val Ala Leu Gly Ile Lys Gly Ala Leu Glu Lys Leu Gly Asp
 290 295 300
 Lys Ala Val Lys Pro Leu Val Val Arg Leu Asp Gly Asn Ala Val Ala
 305 310 315 320
 Glu Gly Arg Lys Ile Leu Glu Glu Phe
 325

<210> 5

<211> 622

<212> DNA

<213> Homo sapiens

<400> 5

acgcgtggcc tagacctgaa tccattctctc acaaaacagt ctccctcctt gccatgggat
 60
 tccctgggggt tttcacaggc cacagctcta atggtctgca gcaggttacc ttgttcccca
 120
 gaacatagct tgtcataaca tctctgcagg gttctcccaa acccctttct gcctggcaac
 180
 agctgacatc acacctagct gtaagtcctt gtagatcgca aattactttt tggagactgg
 240
 gggtagcagg ggcattgggg taatagcctt ctagcccttt ttgagggaaa cacatgggtg
 300

aggcattttt ggggctggga agtggggggc tgggtgcccc tggatggctg tgctggcctc
 360
 tggctgcaag ggagaggggc acaggcaagg acatgacccc cgtcaaccct gagccccctc
 420
 cagaaattta accagagcct gtccctcctt tcttgccctg ccccaacatc tcacaatccc
 480
 tcctgtgatg gcagatgtct ccactctactc tacagacacc tgcaactatc attcccttga
 540
 tccgtggtaa ttaggaggga actcctctgt gaagaaccgc ttctaccatc ctcttttaga
 600
 aactctttct ccactgggat cc
 622

<210> 6
 <211> 121
 <212> PRT
 <213> Homo sapiens

<400> 6
 Met Ser Leu Pro Val Pro Leu Ser Leu Ala Ala Arg Gly Gln His Ser
 1 5 10 15
 His Pro Gly Asp Thr Arg Pro Pro Leu Pro Ser Pro Lys Ile Ala Ser
 20 25 30
 Pro Met Cys Phe Pro Gln Lys Gly Leu Glu Gly Tyr Tyr Pro Asn Ala
 35 40 45
 Pro Ala Thr Pro Ser Leu Gln Lys Val Ile Cys Asp Leu Gln Gly Leu
 50 55 60
 Thr Ala Arg Cys Asp Val Ser Cys Cys Gln Ala Glu Arg Gly Leu Gly
 65 70 75 80
 Glu Pro Cys Arg Asp Val Met Thr Ser Tyr Val Leu Gly Asn Lys Val
 85 90 95
 Thr Cys Cys Arg Pro Leu Glu Leu Trp Pro Val Lys Thr Pro Gly Asn
 100 105 110
 Pro Met Ala Arg Arg Glu Thr Val Leu
 115 120

<210> 7
 <211> 480
 <212> DNA
 <213> Homo sapiens

<400> 7
 agatcttttc agaagcgggtg ggtggggccgt tcatatccac gtgaatgggc ccaaaaggca
 60
 agaaaaccca cacctctcct cccctcccca gtggtgtggt ctttcctagg gacaatagga
 120
 tgaatgggct ttcagtgtgg ggacagcaaa acatgcacta gggcccagag tggcagttct
 180
 cttggtgtgg agagtgcctg ccacaggcct tggccagagc ccgtgagggg gtgctgtgtg
 240
 aaaggccacc tccacgtggg taagcgtgag gacttggaact tctctggcac tgagatggga
 300
 cctcctgcct gtgggagtca tctggccacc accctggggc cagtaaaggt tggagctaga
 360

agggtcgctcc tccctgactt gagctctgag ggctttgcct gccagccag agcggcaagg
 420
 cacaggggac cctcggggac gccatggcc accctgggga agacagggct cctcacgcgt
 480

<210> 8
 <211> 119
 <212> PRT
 <213> Homo sapiens

<400> 8
 Met Gly Phe Gln Cys Gly Asp Ser Lys Thr Cys Thr Arg Ala Gln Ser
 1 5 10 15
 Gly Ser Ser Leu Gly Val Glu Ser Ala Cys His Arg Pro Trp Pro Glu
 20 25 30
 Pro Val Arg Glu Trp Cys Val Lys Gly His Leu His Val Gly Lys Arg
 35 40 45
 Glu Asp Leu Asp Phe Ser Gly Thr Glu Met Gly Pro Pro Ala Cys Gly
 50 55 60
 Ser His Leu Ala Thr Thr Leu Gly Pro Val Lys Val Gly Ala Arg Arg
 65 70 75 80
 Val Val Leu Pro Asp Leu Ser Ser Glu Gly Phe Ala Cys Pro Ala Arg
 85 90 95
 Ala Ala Arg His Arg Gly Pro Ser Gly Thr Pro Met Ala Thr Leu Gly
 100 105 110
 Lys Thr Gly Leu Leu Thr Arg
 115

<210> 9
 <211> 428
 <212> DNA
 <213> Homo sapiens

<400> 9
 tccggaaatt ttccatgaga gactggggca tcgagcagaa gtggatgtct gttctcctgc
 60
 ctctgctgct actttacaat gatccgttct tccccctctc cttcctggtc aacagctggc
 120
 tcccagggat gctggatgac ctctttcagt ccatgttcct gtgcgccctg ctgctcttct
 180
 ggctgtgcgt gtaccacggg attcgtgtcc agggagaaag aaagtgttta actttctatt
 240
 tgctaaatt cttcattgtt ggactattgt ggttggttc tgttacgcta ggaatatggc
 300
 aaacagttaa cgaattacat gatccaatgt accagtatcg agttgatacc ggaaattttc
 360
 agggaatgaa ggtcttcttc atggtggtgg cagcgggtga cattctgtac ctcttgttct
 420
 tgatagtg
 428

<210> 10
 <211> 138
 <212> PRT

<213> Homo sapiens

<400> 10

Met Arg Asp Trp Gly Ile Glu Gln Lys Trp Met Ser Val Leu Leu Pro
 1 5 10 15
 Leu Leu Leu Leu Tyr Asn Asp Pro Phe Phe Pro Leu Ser Phe Leu Val
 20 25 30
 Asn Ser Trp Leu Pro Gly Met Leu Asp Asp Leu Phe Gln Ser Met Phe
 35 40 45
 Leu Cys Ala Leu Leu Leu Phe Trp Leu Cys Val Tyr His Gly Ile Arg
 50 55 60
 Val Gln Gly Glu Arg Lys Cys Leu Thr Phe Tyr Leu Pro Lys Phe Phe
 65 70 75 80
 Ile Val Gly Leu Leu Trp Leu Ala Ser Val Thr Leu Gly Ile Trp Gln
 85 90 95
 Thr Val Asn Glu Leu His Asp Pro Met Tyr Gln Tyr Arg Val Asp Thr
 100 105 110
 Gly Asn Phe Gln Gly Met Lys Val Phe Phe Met Val Val Ala Ala Val
 115 120 125
 Tyr Ile Leu Tyr Leu Leu Phe Leu Ile Val
 130 135

<210> 11

<211> 453

<212> DNA

<213> Homo sapiens

<400> 11

cttaagaatc gcctcactca acgggtcagct tgccgaccat gcccgcctga taatgccccg
 60
 gaatgttgca ggcaaaactca agaccggtgg ccttggtgaa ggtccagggtc agctcggcgg
 120
 acttgccccg ctgcaccagc acgctgttgg ggtcggtcatg cttcatgccg cccatatcgc
 180
 catgccccat ggccggcgtgg tccatcttgc ccatgccggg ggccgtgagc atgccgctgg
 240
 cttgcatctt gagcatttct ttctggtggt cggcgtgcat cgccgcatca cccagattga
 300
 attcgtgcag taactggcct ttgttgacca gcacaaagcg cacggtctca ccggctttta
 360
 catccagagc cttggggcgaa aaggaaatgt cctgcagggt gacttccacg gtgcgcgtgg
 420
 ctttatcggc cgggtgccggg tggccaaacg cgt
 453

<210> 12

<211> 130

<212> PRT

<213> Homo sapiens

<400> 12

Met Leu Gln Ala Asn Ser Arg Pro Val Ala Leu Val Lys Val Gln Val
 1 5 10 15
 Ser Ser Ala Asp Leu Pro Gly Ser Thr Ser Thr Leu Leu Gly Ser Ser

20 25 30
 Cys Phe Met Pro Pro Ile Ser Pro Cys Pro Met Ala Ala Trp Ser Ile
 35 40 45
 Leu Pro Met Pro Val Ala Val Ser Met Pro Leu Ala Cys Ile Leu Ser
 50 55 60
 Ile Ser Phe Trp Cys Ser Ala Cys Ile Ala Ala Ser Pro Arg Leu Asn
 65 70 75 80
 Ser Cys Ser Asn Trp Pro Leu Leu Thr Ser Thr Lys Arg Thr Val Ser
 85 90 95
 Pro Ala Phe Thr Ser Arg Ala Leu Gly Glu Lys Glu Met Ser Cys Arg
 100 105 110
 Val Thr Ser Thr Val Arg Val Ala Leu Ser Ala Gly Ala Gly Trp Pro
 115 120 125
 Asn Ala
 130

<210> 13
 <211> 2034
 <212> DNA
 <213> Homo sapiens

<400> 13
 nacgcgttcg gcgtagtccc cttcctgcc a tccgtcctga gctccctgct gcccgtgctg
 60
 ggcgtggcca agcaggacac ggtgcgcgtg gccttctgct cgggggacct gcggctcctc
 120
 tgggccctac catgctggca ttttcctcca tgtgtcaaac acatgggttc agccagcgaa
 180
 gattccatgg gacctcctcg tgtgggacgt gtgctcccca ccacaaatgg aacgttcctc
 240
 gtttgcattt ggaggggttg gtggtcctgc tggctggagc agcctggggc cagaggaagc
 300
 cgtatcaacc ggctctgcag cgcttcagcg aggggtgcct ggagtaccta gccaacctgg
 360
 accgagcccc agacccacg gtcaggaagg acgcctttgc caccgacatc ttcagcgcct
 420
 acgatgttct cttccatcag tggctgcaga gtcgagaagc caagctccgt cttgccgtgg
 480
 tggaggctct ggggcctatg agccatctgc tgcccagtga gaggtcggaa gagcagctgc
 540
 ccaagctcct ccctgggatt ctgcctctct acaagaagca cgagagacc ttctacttgt
 600
 ccaagagcct gggccagatc ctcgaggcag ctgtgagtgt gggcagccgc aactgggaga
 660
 cccagctgga tgccctcttg gctgcactgc actcccagat ctgtgtgcct gtggagtcct
 720
 caagccccct ggtgatgagt aaccagaagg aggtgctgcg ctgcttcact gtgctggcct
 780
 gcagctcgcc tgaccgcta ctggccttcc tgctgccag gctggacacc accaatgaga
 840
 ggacccgcgt gggcaccctg cagggggcca aacatgtcat caactcaact gctgctcaaa
 900
 tggaagataa aaagcccttt atcctgtctt ccatgaggct tcctctcctg aacaccaaca
 960

gcaaggtgaa gcgggcagtg gtgcaggtga ttagcgccca tggcccacca cggctacctg
 1020
 gagcagcctg gaggtgaggc gatgatcgag taaatcgtgc agcagtgcgc gctgcccccc
 1080
 gagcaggagc ctgagaagcc agggccccggc agcaaggacc ccaaggccga cagcgtgcgg
 1140
 gccatcagcg tgcgcaccct ctacctggtc agcaccaccg tggacaggat gagtcacgtc
 1200
 ctctggccat acctgctcca gtccctcacc cctgtgcgct tcaactggggc cctgactccg
 1260
 ctctgcagga gcctcgtgca tctggcgcag aagaggcagg agggccggggc cgacgccttc
 1320
 ctcatccagt acgacgcccc tgcgagcctc ccgtctccct atgctgtaac cggaagactg
 1380
 ttggttgtgt cttccagccc ctacctaggg gacggacgtg gggcagcggc gctgcgcctc
 1440
 ctcagtgttc tgcacccaaa cattcacctt ttgctgggtc agcattggga aacgactgtc
 1500
 ccgctgctgc tgggggtacct ggatgagcac acagaagaga ccctgccaca ggaggagtgg
 1560
 gaggagaagc tgttgatggt gagggccggg gtacggccca tcctgggcct taagggtgtg
 1620
 tctggcctgg ggggtgctgg ggtggcagag gctgggccac ctgcctcgac ctcacctcgt
 1680
 ggtttggtg gggagccaag gatcaggcag catcaaggct gaagacccca gcagccttgc
 1740
 agcggggggc ttgctgtgac aaggcaccgg ccctctagca gtcgcagccc caagcgtcgg
 1800
 gggcaacctc tcacctgcc tggtagacca actgtggcat ggctgtcccc tgagggttgg
 1860
 ctctgccgcc cccggcctcc gctggaaggc ggtctgcagc ccctgcagcc acagcacatg
 1920
 gggatgtgcc caggctccag ccagccctgt gaggggtcgg gctcccagcc cctcagtggc
 1980
 atcttggcct gcagttcctg cgagacaccc tggccatcat ttctgacaac gcgt
 2034

<210> 14

<211> 222

<212> PRT

<213> Homo sapiens

<400> 14

Ile	Val	Gln	Gln	Cys	Ala	Leu	Pro	Pro	Glu	Gln	Glu	Pro	Glu	Lys	Pro
1				5					10					15	
Gly	Pro	Gly	Ser	Lys	Asp	Pro	Lys	Ala	Asp	Ser	Val	Arg	Ala	Ile	Ser
			20					25				30			
Val	Arg	Thr	Leu	Tyr	Leu	Val	Ser	Thr	Thr	Val	Asp	Arg	Met	Ser	His
		35					40				45				
Val	Leu	Trp	Pro	Tyr	Leu	Leu	Gln	Phe	Leu	Thr	Pro	Val	Arg	Phe	Thr
	50					55				60					
Gly	Ala	Leu	Thr	Pro	Leu	Cys	Arg	Ser	Leu	Val	His	Leu	Ala	Gln	Lys
65					70					75				80	
Arg	Gln	Glu	Ala	Gly	Ala	Asp	Ala	Phe	Leu	Ile	Gln	Tyr	Asp	Ala	His

```

      85              90              95
Ala Ser Leu Pro Ser Pro Tyr Ala Val Thr Gly Arg Leu Leu Val Val
      100              105              110
Ser Ser Ser Pro Tyr Leu Gly Asp Gly Arg Gly Ala Ala Ala Leu Arg
      115              120              125
Leu Leu Ser Val Leu His Pro Asn Ile His Pro Leu Leu Gly Gln His
      130              135              140
Trp Glu Thr Thr Val Pro Leu Leu Leu Gly Tyr Leu Asp Glu His Thr
      145              150              155
Glu Glu Thr Leu Pro Gln Glu Glu Trp Glu Glu Lys Leu Leu Met Val
      165              170              175
Arg Ala Gly Val Arg Pro Ile Leu Gly Leu Lys Val Leu Ser Gly Leu
      180              185              190
Gly Gly Ala Gly Val Ala Glu Ala Gly Pro Pro Ala Ser Thr Ser Pro
      195              200              205
Arg Gly Leu Ala Gly Glu Pro Arg Ile Arg Gln His Gln Gly
      210              215              220

```

<210> 15

<211> 363

<212> DNA

<213> Homo sapiens

<400> 15

```

naccgcttgc tggctcgcca cggcaagggc catgtcggct gcgatatctg caagccggcg
60
gtgggttcga tccttgccctc gtgctggaac cagccgatca tggaccgggc gttggtgccg
120
ttgcaggaca ccaatgacac cttcatggcc aacatgcaga agaacggtac ctattcgatc
180
atcccgcgta tcgccggcgg cgagatcacc ccggacaaac tgatcgccct cggcgcggtg
240
gcgaagaaat acgatctgta caccaagatc accggcggcc agcggatcga cctgttcggc
300
gccagttgc acgaattgcc gcagatctgg ggcgagctgg tggatgccgg attcgagacc
360
ggt
363

```

<210> 16

<211> 121

<212> PRT

<213> Homo sapiens

<400> 16

```

Xaa Ala Leu Leu Ala Arg His Gly Lys Gly His Val Gly Cys Asp Ile
  1              5              10              15
Cys Lys Pro Ala Val Gly Ser Ile Leu Ala Ser Cys Trp Asn Gln Pro
      20              25              30
Ile Met Asp Pro Ala Leu Val Pro Leu Gln Asp Thr Asn Asp Thr Phe
      35              40              45
Met Ala Asn Met Gln Lys Asn Gly Thr Tyr Ser Ile Ile Pro Arg Ile
      50              55              60
Ala Gly Gly Glu Ile Thr Pro Asp Lys Leu Ile Ala Leu Gly Ala Val

```

430

```

65              70              75              80
Phe Pro Thr Pro Glu Ile Pro His Ser Leu Ala Pro Gly Thr Val Ala
              85              90              95
Ile Ser Lys Pro Trp Phe Pro Ala Val Ser Gln Ile Ala Arg
              100              105              110

```

<210> 19
 <211> 515
 <212> DNA
 <213> Homo sapiens

```

<400> 19
cttggtggc agacatgga cctgctccc tcttacccc cagtcttggc aaggatcatg
60
cccccatctc aactatgtta gccagtctgg ctgttcactt agtcactaca gtttgettct
120
cgtctgcagt gcagtcttgg gctataagaa aactggtggc actcaatacc tcccccttt
180
tggcccttct cctcctctgg tccatgggtg gggttggggg gagcccagtt tcagcaccag
240
cagctggagc ccataccaca ctcatttttc agttctggct gtgggagccc ctcccacagg
300
tttcagtcc ccaagcccca ggcctgagtt tttttattg caaaagctgg ttgttgtgtg
360
ggctagctcc caggcgtgtg aggtgcagct tgctaagtaa gagctaggaa agagaatagg
420
gtcctgctgt aggtgtccag tctgaaggaa tgctgggat acttcctcaa gcagttcctt
480
ctcacagtct cctggctgct ccgcatgtca gatct
515

```

<210> 20
 <211> 130
 <212> PRT
 <213> Homo sapiens

```

<400> 20
Met Gly Pro Ala Ser Leu Leu His Pro Ser Leu Gly Lys Asp His Ala
1              5              10              15
Pro Ile Ser Thr Met Leu Ala Ser Leu Ala Val His Leu Val Thr Thr
              20              25              30
Val Cys Phe Ser Ser Ala Val Gln Ser Trp Ala Ile Arg Asn Thr Gly
              35              40              45
Pro Leu Asn Thr Ser Pro Leu Leu Ala Leu Leu Leu Leu Trp Ser Met
              50              55              60
Gly Gly Val Gly Gly Ser Pro Val Ser Ala Pro Ala Ala Gly Ala His
65              70              75              80
Thr Thr Leu Ile Phe Gln Phe Trp Leu Trp Glu Pro Leu Pro Gln Val
              85              90              95
Ser Val Pro Gln Ala Pro Gly Leu Ser Phe Phe Tyr Cys Lys Ser Trp
              100              105              110
Leu Leu Leu Trp Leu Ala Pro Arg Val Arg Cys Ser Leu Leu Ser
              115              120              125
Lys Ser

```

130

<210> 21
 <211> 390
 <212> DNA
 <213> Homo sapiens

<400> 21
 gtgcgcacaa aagagcacgt tcgcaagggg aggaagagcg tgccaccggt tctgccgagc
 60
 tagacgcggg gcctatggtt gcggaggacc atggagtga gcgagtaaga ctagatgatg
 120
 caacaaatgt gcctgagggg gaaatggcac gagccagtgc caatgagggc atgacacctg
 180
 ttaaccacga caaataccct tctgtccttt taaatgaagc ggcccagget tcattactgg
 240
 atacaatgac tgcttgact gatgggttca caattgagca attggagctt acacgatctc
 300
 tatgttatga aagagtatta gcacatcgat cctcatggga tegtccagcc ctgggtcaag
 360
 aattaaagca agttgtccaa ggcattccatn
 390

<210> 22
 <211> 105
 <212> PRT
 <213> Homo sapiens

<400> 22
 Met Val Ala Glu Asp His Gly Val Lys Arg Val Arg Leu Asp Asp Ala
 1 5 10 15
 Thr Asn Val Pro Glu Gly Glu Met Ala Arg Ala Ser Ala Asn Glu Gly
 20 25 30
 Met Thr Pro Val Asn His Asp Lys Tyr Pro Ser Val Leu Leu Asn Glu
 35 40 45
 Ala Ala Gln Ala Ser Leu Leu Asp Thr Met Thr Ala Cys Thr Asp Gly
 50 55 60
 Phe Thr Ile Glu Gln Leu Glu Leu Thr Arg Ser Leu Cys Tyr Glu Arg
 65 70 75 80
 Val Leu Ala His Arg Ser Ser Trp Asp Arg Ser Ala Leu Ala Gln Glu
 85 90 95
 Leu Lys Gln Val Val Gln Gly Ile His
 100 105

<210> 23
 <211> 385
 <212> DNA
 <213> Homo sapiens

<400> 23
 ntctcggagg ccgacagcct ggcgggctgg aagccctcgg tgtaccacgt gctgtctatc
 60
 ctgggcctgt tcgccgtgct gctgtcctgc tgcgcctcgg ccatgtacac cagcgtggag
 120

ggctgggact acgtggactc gctctacttc tgccttcgtca ccttcagcac catcggttc
 180
 ggggacctgg tgagcagcca gcacgccgcc taccggaacc aggggtctta ccgcctgggc
 240
 aacttctctt tcattctgct cggcgtgtgc tgcatttact cgtctttcaa cgctcatctc
 300
 atcctcatca agcaggtgct caactggatg ctgcgcaagc tgagctgccg ctgctgcgcg
 360
 cgctgctgcc cggctcctgg cgcgc
 385

<210> 24
 <211> 128
 <212> PRT
 <213> Homo sapiens

<400> 24
 Xaa Ser Glu Ala Asp Ser Leu Ala Gly Trp Lys Pro Ser Val Tyr His
 1 5 10 15
 Val Leu Leu Ile Leu Gly Leu Phe Ala Val Leu Leu Ser Cys Cys Ala
 20 25 30
 Ser Ala Met Tyr Thr Ser Val Glu Gly Trp Asp Tyr Val Asp Ser Leu
 35 40 45
 Tyr Phe Cys Phe Val Thr Phe Ser Thr Ile Gly Phe Gly Asp Leu Val
 50 55 60
 Ser Ser Gln His Ala Ala Tyr Arg Asn Gln Gly Leu Tyr Arg Leu Gly
 65 70 75 80
 Asn Phe Leu Phe Ile Leu Leu Gly Val Cys Cys Ile Tyr Ser Leu Phe
 85 90 95
 Asn Val Ile Ser Ile Leu Ile Lys Gln Val Leu Asn Trp Met Leu Arg
 100 105 110
 Lys Leu Ser Cys Arg Cys Cys Ala Arg Cys Cys Pro Ala Pro Gly Ala
 115 120 125

<210> 25
 <211> 337
 <212> DNA
 <213> Homo sapiens

<400> 25
 ccatgggaga gaccgtgcat tttcttctag gtctgcgtgg gaagtcactg cagagtttcg
 60
 aggaggggag ttcccagctc tgtatttttg aagggtcagt cttgttgctt ggaccagtga
 120
 ggagccccgt gggatccaga ctcgagtggg tggagccggg gcaggtggga gcagagacac
 180
 tggaggaaag ctggtcgaat gcactgtgta tttggaggca gaaccagcag agggtcctct
 240
 ggggtgagtg tagggcaaaa gagaaagaag gcaccaagcc tggggtctgg gttttctctc
 300
 ttacacttgc tgggtggacg gtggtgccac tgaatga
 337

<210> 26

<211> 111
 <212> PRT
 <213> Homo sapiens

<400> 26
 Met Gly Glu Thr Val His Phe Leu Leu Gly Leu Arg Gly Lys Ser Leu
 1 5 10 15
 Gln Ser Phe Glu Glu Gly Ser Ser Gln Leu Cys Ile Phe Glu Gly Ser
 20 25 30
 Val Leu Leu Leu Gly Pro Val Arg Ser Pro Val Gly Ser Arg Leu Glu
 35 40 45
 Trp Val Glu Pro Gly Gln Val Gly Ala Glu Thr Leu Glu Glu Ser Trp
 50 55 60
 Ser Asn Ala Leu Cys Ile Trp Arg Gln Asn Gln Gln Arg Val Leu Trp
 65 70 75 80
 Val Glu Cys Arg Ala Lys Glu Lys Glu Gly Thr Lys Pro Gly Val Trp
 85 90 95
 Val Phe Ser Leu Thr Leu Ala Gly Trp Thr Val Val Pro Leu Asn
 100 105 110

<210> 27
 <211> 333
 <212> DNA
 <213> Homo sapiens

<400> 27
 ccgacgtcga atatccatgc agccgcgccg aggatggaga gagcgatgga gcaactcaac
 60
 cgcctgacgc gctcgtgctg cgcgcgcgc accgtggagt tgcccaggga taatgaaact
 120
 gctgtttata cattaatgcc aatggttatg gctgatcaac acaggtctgt ttctgaacta
 180
 ctatcaaatt caaaatttga tgtcaattat gcattcggac gtgtgaaaag aagcttgctt
 240
 cacattgcag caaattgtgg atcgggtggaa tgcttggttt tgctgttaaa gaaaggagca
 300
 aatcctaact atcaagatat ttcaggctgt aca
 333

<210> 28
 <211> 111
 <212> PRT
 <213> Homo sapiens

<400> 28
 Pro Thr Ser Asn Ile His Ala Ala Ala Pro Arg Met Glu Arg Ala Met
 1 5 10 15
 Glu Gln Leu Asn Arg Leu Thr Arg Ser Leu Arg Arg Ala Arg Thr Val
 20 25 30
 Glu Leu Pro Glu Asp Asn Glu Thr Ala Val Tyr Thr Leu Met Pro Met
 35 40 45
 Val Met Ala Asp Gln His Arg Ser Val Ser Glu Leu Leu Ser Asn Ser
 50 55 60
 Lys Phe Asp Val Asn Tyr Ala Phe Gly Arg Val Lys Arg Ser Leu Leu

```

65              70              75              80
His Ile Ala Ala Asn Cys Gly Ser Val Glu Cys Leu Val Leu Leu Leu
              85              90              95
Lys Lys Gly Ala Asn Pro Asn Tyr Gln Asp Ile Ser Gly Cys Thr
              100              105              110

```

<210> 29
 <211> 375
 <212> DNA
 <213> Homo sapiens

```

<400> 29
ncgccgtccg tgctggctat tatgacggcg ggtagcgacc agggcgagga ggtcaactcg
60
gagagctatt tgagcgcgtg gacgccgctg agtcccaaag agattcgtca gctgccccgc
120
tacaatatca cgatcaagcg cgtcgtgaac atgacgggca agggccgcac gccgagctgg
180
tactcgctcg tcgtggctgg caatggtcgg ggcctcgtgg gctatggcga aggcaaagat
240
actaacatca gccgcgcgaa caaaaaggcg ttccacgccg cggtgaaaaa catggacttg
300
gtatcgggcc accggtcgaa gattggcgcc aacacgctcg agccccccgt cgagggcccg
360
tggggcgcta cgcgt
375

```

<210> 30
 <211> 125
 <212> PRT
 <213> Homo sapiens

```

<400> 30
Xaa Pro Ser Val Leu Ala Ile Met Thr Ala Gly Ser Asp Gln Gly Glu
1      5      10      15
Glu Val Asn Ser Glu Ser Tyr Leu Ser Ala Val Thr Pro Leu Ser Pro
      20      25      30
Lys Glu Ile Arg Gln Leu Pro Arg Tyr Asn Ile Thr Ile Lys Arg Val
      35      40      45
Val Asn Met Thr Gly Lys Gly Arg Thr Pro Ser Trp Tyr Ser Leu Val
      50      55      60
Val Ala Gly Asn Gly Arg Gly Leu Val Gly Tyr Gly Glu Gly Lys Asp
65      70      75      80
Thr Asn Ile Ser Arg Ala Asn Lys Lys Ala Phe His Ala Ala Val Lys
      85      90      95
Asn Met Asp Leu Val Ser Val His Arg Ser Lys Ser Gly Ala Asn Thr
      100      105      110
Leu Glu Pro Pro Val Glu Gly Arg Trp Gly Ala Thr Arg
      115      120      125

```

<210> 31
 <211> 375
 <212> DNA
 <213> Homo sapiens

<400> 31

accggtcttg gcctcagctt tgctctgaaa ttgaagtcgg tgccaaaagt ggggaagagc
60
gggagcaggc acttacgagc ctgcgcgtca gggatgcttc ctgggcccct gagagtgcag
120
agattcctgg atccagagct gcggctgggc ggctgcagct gcgcctggga gtgcagggct
180
cccgcctgc cagctcaaaa ggaaatgggg gctcctgcct gttcctggct cctgttggcc
240
ctgcagagtg cacaaacctt gccgcgcttc ctccactgca gcttacgtct ttgcagcagc
300
cactcccgat gggctgccac tgccatctgt gagaccataa tgtgtgcaat ttgagactca
360
tggcctgcat tggtt
375

<210> 32

<211> 118

<212> PRT

<213> Homo sapiens

<400> 32

Met	Gln	Ala	Met	Ser	Leu	Lys	Leu	His	Thr	Leu	Trp	Ser	His	Arg	Trp
1				5					10					15	
Gln	Trp	Gln	Pro	Ile	Gly	Ser	Gly	Cys	Cys	Lys	Asp	Val	Ser	Cys	Ser
			20					25					30		
Gly	Gly	Ser	Ala	Ala	Arg	Phe	Val	His	Ser	Ala	Gly	Pro	Thr	Gly	Ala
		35				40						45			
Arg	Asn	Arg	Gln	Glu	Pro	Pro	Phe	Pro	Phe	Glu	Leu	Ala	Gly	Arg	Glu
	50					55					60				
Pro	Cys	Thr	Pro	Arg	Arg	Ser	Cys	Ser	Arg	Pro	Ala	Ala	Ala	Leu	Asp
65				70					75					80	
Pro	Gly	Ile	Ser	Ala	Leu	Ser	Gly	Ala	Gln	Glu	Ala	Ser	Leu	Thr	Arg
			85					90					95		
Arg	Leu	Val	Ser	Ala	Cys	Ser	Arg	Ser	Ser	Pro	Leu	Leu	Ala	Pro	Thr
			100					105					110		
Ser	Ile	Ser	Glu	Gln	Ser										
			115												

<210> 33

<211> 351

<212> DNA

<213> Homo sapiens

<400> 33

ccatgcagcc caaccgttgg cgataaagtc cgttttaggcg ataccaatatt atgggcaacc
60
attgaacaag atttattaac caaaggatgat gagggtgaaat ttgggtggcgg taaaagtgtg
120
cgtgatggta tggcgcaaag cggcaccgca actcgcgaca atccaaatgt attggatttt
180
gtgattacca atgtgatgat cattgatgcc aaattaggca ttatcaaagc cgatattggg
240

attcgcgatg gtcgtattgt cggatatcgga caagcaggta accctgacac catggatgac
 300
 gtcacgccaa acatgattat cgggtgctagc acagaagtac ataacgggtgc a
 351

<210> 34
 <211> 117
 <212> PRT
 <213> Homo sapiens

<400> 34
 Pro Cys Ser Pro Thr Val Gly Asp Lys Val Arg Leu Gly Asp Thr Asn
 1 5 10 15
 Leu Trp Ala Thr Ile Glu Gln Asp Leu Leu Thr Lys Gly Asp Glu Cys
 20 25 30
 Lys Phe Gly Gly Gly Lys Ser Val Arg Asp Gly Met Ala Gln Ser Gly
 35 40 45
 Thr Ala Thr Arg Asp Asn Pro Asn Val Leu Asp Phe Val Ile Thr Asn
 50 55 60
 Val Met Ile Ile Asp Ala Lys Leu Gly Ile Ile Lys Ala Asp Ile Gly
 65 70 75 80
 Ile Arg Asp Gly Arg Ile Val Gly Ile Gly Gln Ala Gly Asn Pro Asp
 85 90 95
 Thr Met Asp Asp Val Thr Pro Asn Met Ile Ile Gly Ala Ser Thr Glu
 100 105 110
 Val His Asn Gly Ala
 115

<210> 35
 <211> 355
 <212> DNA
 <213> Homo sapiens

<400> 35
 nngctagctg caccaccacc tgttcattgca ggcagagcgg ccacccctca tggaagaaga
 60
 ggaatccact gtattgggca caggcttcct gctggacctt ggcaagcagg tgcttggtg
 120
 gtaccaggaa gtccagcgtg tacctcagtg cgtcctcccg ataagtctc tccaccacct
 180
 ggaacacctg gcccaacagg gtgggggctg ttgcctcaaa ggggtggatac agggcggcga
 240
 gagtgctctg cacacagtcc tccactggct caggctccat ggctcggcgc cggggccgcgt
 300
 ccgacgcttg gtcgggcggg cggggccggg cgcgccaccg cctcccttca cgcgt
 355

<210> 36
 <211> 118
 <212> PRT
 <213> Homo sapiens

<400> 36
 Xaa Leu Ala Ala Pro Pro Pro Val His Ala Gly Arg Ala Ala Thr Pro

```

      1           5           10           15
His Gly Arg Arg Gly Ile His Cys Ile Gly His Arg Leu Pro Ala Gly
      20           25           30
Pro Trp Gln Ala Gly Ala Trp Leu Val Pro Gly Ser Pro Ala Cys Thr
      35           40           45
Ser Val Arg Pro Pro Asp Lys Ser Ser Pro Pro Pro Gly Thr Pro Gly
      50           55           60
Pro Thr Gly Trp Gly Leu Leu Pro Gln Arg Val Asp Thr Gly Arg Arg
      65           70           75           80
Glu Cys Ser Ala His Ser Pro Pro Leu Ala Gln Ala Pro Trp Leu Gly
      85           90           95
Ala Gly Pro Arg Pro Thr Leu Gly Arg Ala Gly Gly Ala Gly Arg Ala
      100          105          110
Thr Ala Ser Leu His Ala
      115

```

<210> 37

<211> 492

<212> DNA

<213> Homo sapiens

<400> 37

```

acgcgtggcc ttcgtctgcc accaggaccg actcagcccc accgggtttc cggaccggcc
60
gcaaccatga caagggcgat gttgtgatct gggtaggatc cttctccgac atgctcgagg
120
gatcggatct ctccggcgta gtcacgggtg ttgccgaggc cggctatcgc ccacgggtcc
180
tcgccgacga cgtctgctgc gggttgacgt ggatcactac cggtcagctc gacggtgctc
240
ggcgtcggct gcgcgtgggt ctccagctgc tggcaccctc gtcagacgcc agcgtccag
300
tcgttgggct agagccgtcc tgcactaccg tctggcgtga tgacgcactc cgcctcctgc
360
cagatgatcc gcgcgtccac cgggtagcca gaaacatgca taccgtcgcc gagatgcttg
420
aggcagcaca gtggacccca ccctcgctag caggccacac cctcgtcgct cagccccatt
480
gtcatcccgc gg
492

```

<210> 38

<211> 127

<212> PRT

<213> Homo sapiens

<400> 38

```

Met Leu Glu Gly Ser Asp Leu Ser Ala Val Val Thr Val Leu Ala Glu
  1           5           10           15
Ala Gly Tyr Arg Pro Arg Val Leu Ala Asp Asp Val Cys Cys Gly Leu
      20           25           30
Thr Trp Ile Thr Thr Gly Gln Leu Asp Gly Ala Arg Arg Arg Leu Arg
      35           40           45
Ala Gly Leu Asp Val Leu Ala Pro Leu Ser Asp Ala Ser Val Pro Val

```

50		55		60
Val Gly Leu Glu Pro Ser Cys Thr Thr Val Trp Arg Asp Asp Ala Leu				
65	70	75	80	
Arg Leu Leu Pro Asp Asp Pro Arg Val His Arg Val Ala Arg Asn Met				
	85	90	95	
His Thr Val Ala Glu Met Leu Glu Ala Ala Gln Trp Thr Pro Pro Ser				
	100	105	110	
Leu Ala Gly His Thr Leu Val Ala Gln Pro His Cys His Pro Ala				
115	120	125		

<210> 39

<211> 412

<212> DNA

<213> Homo sapiens

<400> 39

aacgaaggtn ccgtagcgcg tctgaaagcc ctgctgtaaag agcggttccga tcgccgggaa
 60
 gtgatgngca ccgcaaaaat gcaggtgggc gaagccgcga gtccaggcaa gattgtcttt
 120
 gaaatggaag acgtttatta cagcattgcc ggaaaacaac tggtagagcaa cttctctgcg
 180
 caagtcatgc gtggtgataa aattgctgtg attggcccga acggttgtgg taaaacgacg
 240
 ttgctgaaac tgatgttaag taagattcag gcagacagcg gccgtgttca ctgcggtact
 300
 aaactggaag ttgcgtactt cgaccagcac cgtgctgagc tggatcctga gcgtacgggtg
 360
 atggataacc tggccgaagg taagcaggaa gtgatggtaa atggccgtgt an
 412

<210> 40

<211> 137

<212> PRT

<213> Homo sapiens

<400> 40

Asn Glu Gly Xaa Val Arg Ala Leu Lys Ala Leu Arg Lys Glu Arg Ser				
1	5	10	15	
Asp Arg Arg Glu Val Met Xaa Thr Ala Lys Met Gln Val Val Glu Ala				
	20	25	30	
Ala Ser Ser Gly Lys Ile Val Phe Glu Met Glu Asp Val Tyr Tyr Ser				
	35	40	45	
Ile Ala Gly Lys Gln Leu Val Ser Asn Phe Ser Ala Gln Val Met Arg				
	50	55	60	
Gly Asp Lys Ile Ala Leu Ile Gly Pro Asn Gly Cys Gly Lys Thr Thr				
65	70	75	80	
Leu Leu Lys Leu Met Leu Ser Lys Ile Gln Ala Asp Ser Gly Arg Val				
	85	90	95	
His Cys Gly Thr Lys Leu Glu Val Ala Tyr Phe Asp Gln His Arg Ala				
	100	105	110	
Glu Leu Asp Pro Glu Arg Thr Val Met Asp Asn Leu Ala Glu Gly Lys				
	115	120	125	
Gln Glu Val Met Val Asn Gly Arg Val				

130

135

<210> 41
 <211> 1080
 <212> DNA
 <213> Homo sapiens

<400> 41
 gaattcaagt ggacacaggc tccacgcccg cgtctcaccg ataagagcta caagcacaac
 60
 tactatgacg agcgggttcc gctcgaagag cgtcttgagc gcaactgtggc taaggatttc
 120
 gtcacgacgg aggtcgagcc catgtgggat gcggtgatg tcatgcggat gggtaaggat
 180
 ctcttcatcc agcacgggtct gacgacaaat cggaagtcaa tggagtgggt taagcgttac
 240
 taccccgatt tccgcgttca cgcggtgaat ttccctgggg atccgtaccc gatccatc
 300
 gacgcgacct ttgtgccgct tcgtccgggg ctcatcatca acaaccgaa tcgtccactg
 360
 ccgcaggagc agaggaagat ctccgaggcc aatgactggc agatcgttga tgctgctcag
 420
 ccggcgacg acacgcctcc agaattgtgc tactcgtctg tgtggctatc aatgaactgc
 480
 ttggtacttg atccgaagac ggtcatctgc gaggtctcgg aagttcatca gatggagcag
 540
 atggacaagc tgggtatgaa cgtcatcccc gtcgccttcc gtgacgcgta cccattcggt
 600
 ggaggtctcc actgcgccac agctgatgta tatcgccaag gtacctgtga ggactacttc
 660
 ccgaatcagg tcgacgaccc gaccttgggtg tgagaaaacc ccgtggtcac gtcatgactg
 720
 acggatctcg gtggctcggg acggaactta cgttgtccgt taccgggccg ccgggtctga
 780
 tatggcagta tcacgcctag caaaaaggag catgtcatgg acatggagcc gggcatcatc
 840
 aacgtcaaac aggaagtcc aggcgtcggg acgatgaacc agaaagtggg attcgtgtcc
 900
 atgtttcttt ctgcaacggg tatgggggtg gtgggtactt tcgggcgtct cagcactcct
 960
 gtggatccca cgacgggcag taagtacatc atcgggtgatt ttttggccac tggtaggatg
 1020
 atagtcgggg tcctgggatt tctgcttatt atcgtcatatc ttggaaaatg gtctgagctc
 1080

<210> 42
 <211> 230
 <212> PRT
 <213> Homo sapiens

<400> 42
 Glu Phe Lys Trp Thr Gln Ala Pro Arg Pro Arg Leu Thr Asp Lys Ser
 1 5 10 15
 Tyr Lys His Asn Tyr Tyr Asp Glu Arg Val Ser Leu Glu Glu Arg Leu

```
<210> 43
<211> 358
<212> DNA
<213> Homo sapiens
```

```
<210> 44
<211> 105
<212> PRT
<213> Homo sapiens
```

441

1					5					10					15					
Leu	Gly	His	Arg	Gly	Cys	Leu	Pro	Ala	Ala	Pro	Gly	Arg	Ala	Val	Asp					
				20					25					30						
Thr	Thr	Ala	Trp	Ser	Ala	Val	Pro	Gly	His	Cys	His	Ala	His	Arg	Cys					
				35					40					45						
Pro	Leu	Arg	Met	Asp	Leu	Gly	Phe	Arg	Ile	Arg	Val	Ala	Tyr	Gln	Arg					
				50					55					60						
Glu	Ser	Gln	Ile	Leu	Lys	Glu	Val	Gln	Ser	Pro	Glu	Gly	Met	Ile	Ser					
				65					70					75						
Leu	Arg	Asp	Thr	Ala	Ala	Ser	Leu	Arg	Leu	Glu	Arg	Asp	Thr	Arg	Gln					
				85					90					95						
Leu	Pro	Leu	Leu	Thr	Ser	Ala	Leu	His												
				100					105											

<210> 45

<211> 905

<212> DNA

<213> Homo sapiens

<400> 45

```

gtcgacgata aaggagtatt tgcgcagcag cagtatgatg ctctcgttga ggcggggttc
60
gcggctcctg gaatcccaga gcagtatggt ggcgacggtg cggatgcatg tgcgtccgca
120
ataatcatgg aagaggctgc tgcagttctg gcgtcgctcg ccaccgtcat atcgtccaat
180
gagcttggtg ccgtccctct cctcaaatac ggtagcgagg agcagaggaa acgttatctt
240
tctgaagttg cttcgggtaa ggcacttttc ggatatgcgc tctccgaggc tgatgctgga
300
tcagatccag ctgcacttaa gtgtcgagcc gacgaagatg gggacagttt cgtcctgaat
360
ggcggttaagg cttgggtcac ggaggctggc gaggccaagt acctggtgat atttgcggtt
420
actgaccag acgatccgcg ccacagaatc agcgcgttga tgggccatgc agatgaccgc
480
ggcattagct acgggggtcc ggagcacaaa atgggggatac gcgggtcagt taccagggaa
540
gtggttttca agaatacgcg tatccccaag gaacgagtaa ttggccgctg agggcacggt
600
ctgagtgttg ctctaggtac gcttgataac tctcggtgtt cgattgctgc tcaagcagtg
660
ggaattgccc aaggagcttt agacattgcc acggattacg tccagaagcg caagcagttt
720
ggccagccac tgtccaattt tgagggaatc cagttcatgc tcgcagacat ggcaatgcgt
780
ttggaggcgg cgcgagcgct gacatactct gcagctgacg gtagtgggcg ccagactgac
840
gatgtgagtt acttcggcgc ggcgggccaaa tgtttcgctt ccgacacagc gatggcagtg
900
tgcac
905

```

<210> 46

<211> 301
 <212> PRT
 <213> Homo sapiens

<400> 46

```

Val Asp Asp Lys Gly Val Phe Ala Gln Gln Gln Tyr Asp Ala Leu Val
 1           5           10           15
Glu Ala Gly Phe Ala Ala Pro Gly Ile Pro Glu Gln Tyr Gly Gly Asp
           20           25           30
Gly Ala Asp Ala Ile Ala Ser Ala Ile Ile Met Glu Glu Val Ala Arg
           35           40           45
Val Cys Ala Ser Ser Ser Thr Val Ile Ser Ser Asn Glu Leu Gly Thr
 50           55           60
Val Pro Leu Leu Lys Tyr Gly Ser Glu Glu Gln Arg Lys Arg Tyr Leu
 65           70           75           80
Ser Glu Val Ala Ser Gly Lys Ala Leu Phe Gly Tyr Ala Leu Ser Glu
           85           90           95
Ala Asp Ala Gly Ser Asp Pro Ala Ala Leu Lys Cys Arg Ala Asp Glu
           100          105          110
Asp Gly Asp Ser Phe Val Leu Asn Gly Val Lys Ala Trp Val Thr Glu
           115          120          125
Ala Gly Glu Ala Lys Tyr Leu Val Ile Phe Ala Val Thr Asp Pro Asp
 130          135          140
Asp Pro Arg His Arg Ile Ser Ala Leu Met Val His Ala Asp Asp Pro
 145          150          155          160
Gly Ile Ser Tyr Gly Ala Pro Glu His Lys Met Gly Ile Arg Gly Ser
           165          170          175
Val Thr Arg Glu Val Val Phe Lys Asn Thr Arg Ile Pro Lys Glu Arg
           180          185          190
Val Ile Gly Arg Arg Gly His Gly Leu Ser Val Ala Leu Gly Thr Leu
           195          200          205
Asp Asn Ser Arg Val Ser Ile Ala Ala Gln Ala Val Gly Ile Ala Gln
 210          215          220
Gly Ala Leu Asp Ile Ala Thr Asp Tyr Val Gln Lys Arg Lys Gln Phe
 225          230          235          240
Gly Gln Pro Leu Ser Asn Phe Glu Gly Ile Gln Phe Met Leu Ala Asp
           245          250          255
Met Ala Met Arg Leu Glu Ala Ala Arg Ala Leu Thr Tyr Ser Ala Ala
           260          265          270
Asp Arg Ser Gly Arg Gln Thr Asp Asp Val Ser Tyr Phe Gly Ala Ala
           275          280          285
Ala Lys Cys Phe Ala Ser Asp Thr Ala Met Ala Val Cys
 290          295          300

```

<210> 47
 <211> 379
 <212> DNA
 <213> Homo sapiens

<400> 47

```

aagcttgtag agctagtcg aagcggactg tcggtacgcc aagctgctaa aagatgtggg
 60
atgcatttta ccgctgcgta tgccgtagct acggaagctg ggtgccatat ccggttaagt
 120

```


cagtatgctc ggaaagtccg ccagacgcag ttaagagtgg aatacctgcg ccttcggctg
 180
 gcgagcctgc ctggtggtga tgctggcgcg gcagtaggaa ttgatcgctg actgcgttta
 240
 gatttcgaaa aaggactcac caaatcccag ggtcgacgag aagagttcat acccgtcggc
 300
 gaagacgcca gcacgtataa cagacttatg aaagcgctgc gccaacgcca tgatgtcatc
 360
 aaatccggaa agcttgccc
 379

<210> 48
 <211> 106
 <212> PRT
 <213> Homo sapiens

<400> 48
 Met His Leu Thr Ala Ala Tyr Ala Val Ala Thr Glu Ala Gly Cys His
 1 5 10 15
 Ile Arg Leu Ser Gln Tyr Ala Arg Lys Val Arg Gln Thr Gln Leu Arg
 20 25 30
 Val Glu Tyr Leu Arg Leu Arg Leu Ala Ser Leu Pro Gly Gly Asp Ala
 35 40 45
 Gly Ala Ala Val Gly Ile Asp Arg Arg Leu Arg Leu Asp Phe Glu Lys
 50 55 60
 Gly Leu Thr Lys Ser Gln Gly Arg Arg Glu Glu Phe Ile Pro Val Gly
 65 70 75 80
 Glu Asp Ala Ser Thr Tyr Asn Arg Leu Met Lys Ala Leu Arg Gln Arg
 85 90 95
 His Asp Val Ile Lys Ser Gly Lys Leu Ala
 100 105

<210> 49
 <211> 309
 <212> DNA
 <213> Homo sapiens

<400> 49
 tgatcatgat gctggcatgg actattctgg tccctgttcc tctctcacct gctgaaggac
 60
 atccctctaa tttttgtgtc tccttctgta tcatcaaatt ttcctctctt actgagtctc
 120
 ttgcatctcc ttggaagcat gctgtactat gtcccatcct taaagaactc cccttgctcg
 180
 cacattaccc tctgccagct ggctcatttt tctgctcccc ttacagggga aactcttcaa
 240
 aaagttatct ccacctcctt ccattctcatg ttctcttgaa cctgcagtac tgggtgctcc
 300
 ctcccttttg
 309

<210> 50
 <211> 101
 <212> PRT

<213> Homo sapiens

<400> 50

```

Met Met Leu Ala Trp Thr Ile Leu Val Pro Val Pro Leu Ser Pro Ala
 1              5              10              15
Glu Gly His Pro Ser Asn Phe Cys Val Ser Phe Cys Ile Ile Lys Phe
      20              25              30
Ser Leu Ser Thr Glu Ser Leu Ala Ser Pro Trp Lys His Ala Val Leu
      35              40              45
Cys Pro Ile Leu Lys Glu Leu Pro Leu Ser Ala His Tyr Pro Leu Pro
      50              55              60
Ala Gly Ser Phe Phe Cys Ser Pro Leu Gln Gly Asn Ser Ser Lys Ser
65              70              75              80
Tyr Leu His Leu Leu Pro Ser His Val Leu Leu Asn Leu Gln Tyr Trp
      85              90              95
Val Leu Pro Pro Phe
      100

```

<210> 51

<211> 512

<212> DNA

<213> Homo sapiens

<400> 51

```

agatctttga agaattgccca cactgtcttc ctccctgctt ataatttcct tattccctag
60
gatgtgatecc ttgttcttgg ggctcacat ggcagctgga tctctggcga ttgcatctga
120
gttccagaca ccaggatgga aaagaaaaga aggaggggca agaggaaccc ccagatgctc
180
cttaagagct actgcgtggc attcccactt gcatctcatt tgctcgatcg ctgtcactgt
240
gccctaacga gctgcaagga cactggggaa atgagtctgt cttgtacttc atgtgccctt
300
caaaatcttc tgttgctgag ggagaagagg ccagccggta ttgaggaaca actagcactt
360
tctgcttccg cgtcccaggg ggacgtgggt gtgttgaatc cacaccgggg gtgcggacct
420
ctgaggctgg gctggatggg acatcaggtg ggccctctgt ttcatttatg tgacctccca
480
tcaggctctc tggttgatc ctgctttcta ga
512

```

<210> 52

<211> 125

<212> PRT

<213> Homo sapiens

<400> 52

```

Met Glu Lys Lys Arg Arg Gly Lys Arg Asn Pro Gln Met Leu Leu
 1              5              10              15
Lys Ser Tyr Cys Val Ala Phe Pro Leu Ala Ser His Leu Leu Asp Arg
      20              25              30
Cys His Cys Ala Leu Thr Ser Cys Lys Asp Thr Gly Glu Met Ser Leu

```

```

      35          40          45
Ser Cys Thr Ser Cys Ala Pro Gln Asn Leu Leu Leu Leu Arg Glu Lys
      50          55          60
Arg Pro Ala Gly Ile Glu Glu Gln Leu Ala Leu Ser Ala Ser Ala Ser
      65          70          75          80
Gln Gly Asp Val Gly Val Leu Asn Pro His Arg Gly Cys Gly Pro Leu
      85          90          95
Arg Leu Gly Trp Met Gly His Gln Val Gly Pro Leu Phe His Leu Cys
      100          105          110
Asp Leu Pro Ser Gly Leu Leu Val Gly Ser Cys Phe Leu
      115          120          125

```

<210> 53
 <211> 474
 <212> DNA
 <213> Homo sapiens

```

<400> 53
accggtacac ctacgtcacc cgtaaaaacc gacgcaatac ccggatcgcc tcgtcctcaa
60
aaaattcgat cccgtcgtgc gtcgtcacat tgagttcaag gaggcccgtt aatggccaaa
120
aagtccaaga ttgtcgccca gaagaaacgt gagaagctcg tagcccaata cgccgaaagg
180
cgcgccgaac tcaaggccat catgaagtgc ccaactgcct cattggacga acgcatggag
240
gcatcgcgta agctgtctcg cctgccgcgc gattcatccc ccgtgcgggtt acgtaaccgt
300
gaccaagtcg acgggcgtcc ccgcggctac gttggcaagg ccggtgtgtc ccgtatccgt
360
ttccgtgaga tggcccaccc cggcgaactc cccggaatcg cgaagtcaag ctggtgaagc
420
catggcagta ccgaagcgaa agaagtcccg ttcgaccacg cgtcataggc gggc
474

```

<210> 54
 <211> 101
 <212> PRT
 <213> Homo sapiens

```

<400> 54
Met Ala Lys Lys Ser Lys Ile Val Ala Gln Lys Lys Arg Glu Lys Leu
1      5      10      15
Val Ala Gln Tyr Ala Glu Arg Arg Ala Glu Leu Lys Ala Ile Met Lys
      20      25      30
Cys Pro Thr Ala Ser Leu Asp Glu Arg Met Glu Ala Ser Arg Lys Leu
      35      40      45
Ser Arg Leu Pro Arg Asp Ser Ser Pro Val Arg Leu Arg Asn Arg Asp
      50      55      60
Gln Val Asp Gly Arg Pro Arg Gly Tyr Val Gly Lys Ala Gly Val Ser
      65      70      75      80
Arg Ile Arg Phe Arg Glu Met Ala His Arg Gly Glu Leu Pro Gly Ile
      85      90      95
Ala Lys Ser Ser Trp

```

100

<210> 55
 <211> 378
 <212> DNA
 <213> Homo sapiens

<400> 55
 ccatggccca ggacagccgg catatcggct acgactacgg tacaccggtg gcgccacagt
 60
 tcggcgccagc caagcccgca gcgtgctgcc aggcgcaagc gacaaacacc ggcccgtggg
 120
 tgggtgttcga ccatgtgcgt tgcacccacg acacctttct gatcgacgtc tttctcaacc
 180
 agcccgatgc caccgcgcag caggtcaatg ccgacaaccc gcactacgtc gggcggtttca
 240
 gccgcacatcg catgggcctg gtggatgaca agggccggtg cattaccag ggcgtatcgc
 300
 gcgcggtgaa tgcggcgccg agcaccaagg cgctgaacct gggaccgagt gacgcggcgc
 360
 agttatcggg gaggcgta
 378

<210> 56
 <211> 125
 <212> PRT
 <213> Homo sapiens

<400> 56
 Met Ala Gln Asp Ser Arg His Ile Gly Tyr Asp Tyr Gly Thr Pro Val
 1 5 10 15
 Ala Pro Gln Phe Gly Ala Ala Lys Pro Ala Ala Cys Cys Gln Ala Gln
 20 25 30
 Ala Thr Asn Thr Gly Pro Trp Val Val Phe Asp His Val Arg Cys Thr
 35 40 45
 His Asp Thr Phe Leu Ile Asp Val Phe Leu Asn Gln Pro Asp Ala Thr
 50 55 60
 Ala Gln Gln Val Asn Ala Asp Asn Pro His Tyr Val Gly Arg Phe Ser
 65 70 75 80
 Arg Ile Gly Met Gly Leu Val Asp Asp Lys Gly Arg Cys Ile Thr Gln
 85 90 95
 Gly Val Ser Arg Ala Leu Asn Ala Ala Arg Ser Thr Lys Ala Leu Asn
 100 105 110
 Leu Gly Pro Ser Asp Ala Ala Gln Leu Ser Val Arg Arg
 115 120 125

<210> 57
 <211> 388
 <212> DNA
 <213> Homo sapiens

<400> 57
 agaccacccc gacacagatc aggagtcgtc atgtccagaa agaagaaggt cggcatcctc
 60

accgcaggcg gtgattgccc cgggctcaac gccgctatcc gcggatttgg caaggctgcc
 120
 atccgccagc acgacatgga gctcatcggt attcaggacg gctttcttgg attggcggga
 180
 aaccgcacca tctcccttgg cccgcgtgcc ctctcaggca tcttgacggt cggcgggacc
 240
 atcctgggaa ctagccgtga caaggtaaat cacatgatta tcgacggcga ggaacgggat
 300
 atgggtcccca ccaccgtcga gaattacgag aagctggggc ttgacgcttt ggtgactttg
 360
 ggtggcgggtg gcaccgccaa gaacgcgt
 388

<210> 58
 <211> 129
 <212> PRT
 <213> Homo sapiens

<400> 58
 Arg Pro Thr Arg His Arg Ser Gly Val Val Met Ser Arg Lys Lys Lys
 1 5 10 15
 Val Gly Ile Leu Thr Ala Gly Gly Asp Cys Pro Gly Leu Asn Ala Ala
 20 25 30
 Ile Arg Gly Phe Gly Lys Ala Ala Ile Arg Gln His Asp Met Glu Leu
 35 40 45
 Ile Gly Ile Gln Asp Gly Phe Leu Gly Leu Ala Gly Asn Arg Thr Ile
 50 55 60
 Ser Leu Gly Pro Arg Ala Leu Ser Gly Ile Leu Thr Val Gly Gly Thr
 65 70 75 80
 Ile Leu Gly Thr Ser Arg Asp Lys Val Asn His Met Ile Ile Asp Gly
 85 90 95
 Glu Glu Arg Asp Met Val Pro Thr Thr Val Glu Asn Tyr Glu Lys Leu
 100 105 110
 Gly Leu Asp Ala Leu Val Thr Leu Gly Gly Gly Gly Thr Ala Lys Asn
 115 120 125
 Ala

<210> 59
 <211> 417
 <212> DNA
 <213> Homo sapiens

<400> 59
 ggtaccatcg gagctcgaca agaaatgggtt ggggtgaagtc gtggcttctg ctccacccag
 60
 tgcctctcatg ggtcagccca cctgaatatc ttcattgctg tgcattttctc ctgatgttca
 120
 cgtgtgccct gtgtttttac gcatctgtga tcgtgcaccc acgcgtctca gagaggagcc
 180
 cgtttgggaa tccggagaat gtgcgctggc ggaagagcgt cacacactgg aagcaaacct
 240
 cagaccgcgt ggacaagacc aaggatgaaa tggaacacga ggccttggtg gaagggaacc
 300

tggcaaccga ggcaagccta gtggttctgg acacactgga gatcatcgtg cagacggtga
 360
 tgctttcaga agcccgggag agcgtcttgg gggcagtgtc gaaggttgtg ctgtaca
 417

<210> 60
 <211> 101
 <212> PRT
 <213> Homo sapiens

<400> 60
 Met Phe Thr Cys Ala Leu Cys Phe Tyr Ala Ser Val Ile Val His Pro
 1 5 10 15
 Arg Val Ser Glu Arg Ser Pro Phe Gly Asn Pro Glu Asn Val Arg Trp
 20 25 30
 Arg Lys Ser Val Thr His Trp Lys Gln Thr Ser Asp Arg Val Asp Lys
 35 40 45
 Thr Lys Asp Glu Met Glu His Glu Ala Leu Val Glu Gly Asn Leu Ala
 50 55 60
 Thr Glu Ala Ser Leu Val Val Leu Asp Thr Leu Glu Ile Ile Val Gln
 65 70 75 80
 Thr Val Met Leu Ser Glu Ala Arg Glu Ser Val Leu Gly Ala Val Leu
 85 90 95
 Lys Val Val Leu Tyr
 100

<210> 61
 <211> 304
 <212> DNA
 <213> Homo sapiens

<400> 61
 agatcttcac agccttagac ttttttcatg ggtgccttac agttttggag gtccttatcc
 60
 gcacacatat ttgcaggctt ggagagagtg tgtgggggca tgtactttcg gtgggtcaag
 120
 tatgaagaag caggccttat aaacacatat tctgacctta acctgtactt cagaagagga
 180
 ccgctgactc accaaggagg cctgaaggac aaggcagcat ctctgtcttc acatgagtcc
 240
 tcccctagac cgggcccattg gccaggcctg accacagagc tcccattgcc tttcctgcac
 300
 gcgt
 304

<210> 62
 <211> 92
 <212> PRT
 <213> Homo sapiens

<400> 62
 Met Gly Ala Leu Gln Phe Trp Arg Ser Leu Ser Ala His Ile Phe Ala
 1 5 10 15

Gly Leu Glu Arg Val Cys Gly Gly Met Tyr Phe Arg Trp Val Lys Tyr
 20 25 30
 Glu Glu Ala Gly Leu Ile Asn Thr Tyr Ser Asp Leu Asn Leu Tyr Phe
 35 40 45
 Arg Arg Gly Pro Leu Thr His Gln Gly Gly Leu Lys Asp Lys Ala Ala
 50 55 60
 Ser Leu Ser Ser His Glu Ser Ser Pro Arg Pro Gly Pro Trp Pro Gly
 65 70 75 80
 Leu Thr Thr Glu Leu Pro Leu Pro Phe Leu His Ala
 85 90

<210> 63

<211> 577

<212> DNA

<213> Homo sapiens

<400> 63

cgcgtaagg gggctacac cgggacgatt aacgcctcgg tgggagtatt catcaccgcg
 60
 ctgacgggtg tagctgggtg gctcacccta gccgggcgta tcagtgtcgg ggaactcgtc
 120
 accgtggtcg ggctggccca aacctcggc cctccgtgc gagcactggg cgtcgacacc
 180
 gcgacgatgt tggccaccgc ccacgcctcc ggggaccgat tctgtgagtt gcgtgatagc
 240
 cggcagcct ggcagatcca ccccgacgac ggtgcccgca ccacaccggg tgatggcccc
 300
 gtggagtgc acatcccgtt cagggatttc cagcttgacg tcgccggcgg caccatgtg
 360
 ggtatcatgg cgctcaatc ggtctgtgac gccttgccg aggcgataga ccacggctcc
 420
 gagaccgtct tgaatgggtt tcccgccagt cgctcaacc ctgcccacg gcgtcgtctg
 480
 gtgctggtgg ctccccgtc ccccgaaactg ttcgacgata ctgcccggtc gaacatcgtg
 540
 cttgacagcc agacgactgt cgccaggctg aatgcat
 577

<210> 64

<211> 192

<212> PRT

<213> Homo sapiens

<400> 64

Arg Val Lys Gly Val Tyr Thr Gly Thr Ile Asn Ala Ser Val Gly Val
 1 5 10 15
 Phe Ile Thr Ala Leu Thr Val Leu Ala Gly Trp Leu Thr Leu Ala Gly
 20 25 30
 Arg Ile Ser Val Gly Glu Leu Val Thr Val Val Gly Leu Ala Gln Thr
 35 40 45
 Leu Gly Pro Pro Leu Arg Ala Leu Gly Val Asp Thr Ala Thr Met Leu
 50 55 60
 Ala Thr Ala His Ala Ser Gly Asp Arg Phe Cys Glu Leu Arg Asp Ser
 65 70 75 80

```

Pro Ala Ala Trp Gln Ile His Pro Asp Asp Gly Ala Arg Thr Thr Pro
      85          90          95
Gly Asp Gly Pro Val Glu Leu His Ile Pro Val Arg Asp Phe Gln Leu
      100        105        110
Asp Val Ala Gly Gly Thr His Val Gly Ile Met Ala Pro Gln Ser Val
      115        120        125
Cys Asp Ala Leu Ala Glu Ala Ile Asp His Gly Ser Glu Thr Val Leu
      130        135        140
Asn Gly Val Pro Ala Ser Arg Leu Asn Pro Ala Gln Arg Arg Arg Leu
      145        150        155        160
Val Leu Val Ala Pro Arg Ser Pro Glu Leu Phe Asp Asp Thr Ala Arg
      165        170        175
Ala Asn Ile Val Leu Asp Ser Gln Thr Thr Val Ala Arg Leu Asn Ala
      180        185        190

```

<210> 65
 <211> 339
 <212> DNA
 <213> Homo sapiens

```

<400> 65
gtcgaccgcg ccttgggatc gctcgaagg ggcagcctgg accaggtagc ggaagaagtc
60
aagaaggccg ctttcaagat caccgcgcc gggcaactag tgggcacccat ggcctccgag
120
cgcccttgccg tacccttcgg catcatcgac ctttcgcttg cccctactgc cgaattggga
180
gattcgggggg cccacatcct tgagcatatg ggattggacc aagtaggcac gcacggcaca
240
actgctgctt tggctctgct taacgacgcc gtaaagaaag gcggcatgat ggcttgcccc
300
cgcgctggcg gtttgcctgg ctcccttcac cggggctcc
339

```

<210> 66
 <211> 113
 <212> PRT
 <213> Homo sapiens

```

<400> 66
Val Asp Arg Ala Leu Gly Ser Leu Glu Gly Ala Ser Leu Asp Gln Val
1      5      10      15
Ala Glu Glu Val Lys Lys Ala Ala Phe Lys Ile Thr Arg Ala Gly Gln
20     25     30
Leu Val Gly Thr Met Ala Ser Glu Arg Leu Gly Val Pro Phe Gly Ile
35     40     45
Ile Asp Leu Ser Leu Ala Pro Thr Ala Glu Leu Gly Asp Ser Gly Ala
50     55     60
His Ile Leu Glu His Met Gly Leu Asp Gln Val Gly Thr His Gly Thr
65     70     75     80
Thr Ala Ala Leu Ala Leu Leu Asn Asp Ala Val Lys Lys Gly Gly Met
85     90     95
Met Ala Cys Pro Arg Val Gly Gly Leu Ser Gly Ser Phe Ile Pro Gly
100    105    110

```


Ser

<210> 67

<211> 446

<212> DNA

<213> Homo sapiens

<400> 67

tgatcataaa ccacgcgtca ccgaggggat gtggcacacc tacctgcgcg tcgcagatgc
 60
 cgcacaggca cgggtcaggg gcgttcgcgg cgccagctgg cacaacttcg cgaccggtga
 120
 caaggggtcc ttcgacgcca acgagcttgc cgtaactcct gatactgaca ccgtcatcca
 180
 gggagtcggg cccgccctag cctcctcga ttcagcgtgg ggacgccaga tccacgtgga
 240
 gacaacaggg tgtcccagtg ccgtgggtctg gaatccacgc tcctcgtcga cacatgccga
 300
 taacccgaca gccacggcat ggcgcgattt cgtatgcgtc gagaccgggg cctgcaagga
 360
 caatgcggtc attgttgccc cacacagcga cctcaccatg tccacacgga ttagcgtcga
 420
 aacgttgtga tcgctgcatg gatatt
 446

<210> 68

<211> 133

<212> PRT

<213> Homo sapiens

<400> 68

Met Trp His Thr Tyr Leu Arg Val Ala Asp Ala Ala Gln Ala Arg Val
 1 5 10 15
 Arg Gly Val Arg Gly Ala Ser Trp His Asn Phe Ala Thr Gly Asp Lys
 20 25 30
 Gly Ser Phe Asp Ala Asn Glu Leu Ala Val Thr Pro Asp Thr Asp Thr
 35 40 45
 Val Ile Gln Gly Val Gly Pro Ala Leu Ala Leu Leu Asp Ser Ala Trp
 50 55 60
 Gly Arg Gln Ile His Val Glu Thr Thr Gly Cys Pro Ser Ala Val Val
 65 70 75 80
 Trp Asn Pro Arg Ser Ser Ser Thr His Ala Asp Asn Pro Thr Ala Gln
 85 90 95
 Ala Trp Arg Asp Phe Val Cys Val Glu Thr Gly Ala Cys Lys Asp Asn
 100 105 110
 Ala Val Ile Val Ala Pro His Ser Asp Leu Thr Met Ser Thr Arg Ile
 115 120 125
 Ser Val Glu Thr Leu
 130

<210> 69

<211> 552

<212> DNA

<213> Homo sapiens

<400> 69

nnaagggttaa ggagaaaagc aaggaccttg caaagagagc ctctgtgccg gagaggctgg
60
ccctcaagga ggagccaaaa gaagacccca gtggagcagc tgtgcccag atgccaaaaa
120
agtcctccaa gattgccagc ttcaccccca aaggggggaa gctcaacagt gccagaagg
180
agcncatgg ccccttcctt cagtgaata ccaaaaccag gaatgaaaag catgcccggg
240
aaatccccaa gtgccccagc gccttccaag gaaggggagc ggagccggag tgggaagctg
300
agctcaggac tccccagca gaagcccag ctggacggca gacactccag ttcctcttcc
360
agcctggcgt cctcagaagg aaaaggccca ggagggacca cctgaacca cagcatcagc
420
agccagactg tcagtgggtc tgcgggacc acccagacca caggaagcaa tnnaccgtca
480
gtgttcagct acctcagccc cagcagcaat acaaccatcc caacactgcc acggttgac
540
ctttctgta ca
552

<210> 70

<211> 184

<212> PRT

<213> Homo sapiens

<400> 70

Xaa	Arg	Val	Arg	Arg	Lys	Ala	Arg	Thr	Leu	Gln	Arg	Glu	Pro	Leu	Cys
1			5						10					15	
Arg	Arg	Gly	Trp	Pro	Ser	Arg	Arg	Ser	Gln	Lys	Lys	Thr	Pro	Val	Glu
			20					25					30		
Gln	Leu	Cys	Pro	Arg	Cys	Gln	Lys	Ser	Pro	Pro	Arg	Leu	Pro	Ala	Ser
		35					40					45			
Ser	Pro	Lys	Gly	Gly	Ser	Ser	Thr	Val	Pro	Arg	Arg	Ser	Xaa	Met	Ala
	50					55					60				
Pro	Ser	Leu	Ser	Gly	Ile	Pro	Lys	Pro	Gly	Met	Lys	Ser	Met	Pro	Gly
65					70					75				80	
Lys	Ser	Pro	Ser	Ala	Pro	Ala	Pro	Ser	Lys	Glu	Gly	Glu	Arg	Ser	Arg
				85					90					95	
Ser	Gly	Lys	Leu	Ser	Ser	Gly	Leu	Pro	Gln	Gln	Lys	Pro	Gln	Leu	Asp
			100					105					110		
Gly	Arg	His	Ser	Ser	Ser	Ser	Ser	Leu	Ala	Ser	Ser	Glu	Gly	Lys	
		115					120					125			
Gly	Pro	Gly	Gly	Thr	Thr	Leu	Asn	His	Ser	Ile	Ser	Ser	Gln	Thr	Val
	130					135					140				
Ser	Gly	Ser	Val	Gly	Thr	Thr	Gln	Thr	Thr	Gly	Ser	Asn	Xaa	Pro	Ser
145					150					155				160	
Val	Phe	Ser	Tyr	Leu	Ser	Pro	Ser	Ser	Asn	Thr	Thr	Ile	Pro	Thr	Leu
				165					170					175	
Pro	Arg	Leu	His	Leu	Ser	Cys	Thr								
				180											

<210> 71
 <211> 316
 <212> DNA
 <213> Homo sapiens

<400> 71
 cgcggttgaaa tggcggttcga acttaaactg ttacatattg actccgtgcc attaaacatt
 60
 ttgaatcctg ttaaaggagc tccatttgaa agcaacgaag ctttacgtcc tttaaatac
 120
 ttacgtacct tcgccgtatt ccgtttcatc ttgccaaacg cattgatacg aactgcaggt
 180
 ggccgcgaag taaatctacg agacttacaa gcttatgctc taaaagggtg cctaaacggt
 240
 atcatgggtg gtggctactt aactactggc ggctgttcac ctcaagacga tctccaaatg
 300
 attcaagact tggagt
 316

<210> 72
 <211> 105
 <212> PRT
 <213> Homo sapiens

<400> 72
 Arg Val Glu Met Ala Phe Glu Leu Lys Arg Leu His Ile Asp Ser Val
 1 5 10 15
 Pro Leu Asn Ile Leu Asn Pro Val Lys Gly Thr Pro Phe Glu Ser Asn
 20 25 30
 Glu Ala Leu Arg Pro Leu Asn Ile Leu Arg Thr Phe Ala Val Phe Arg
 35 40 45
 Phe Ile Leu Pro Asn Ala Leu Ile Arg Thr Ala Gly Gly Arg Glu Val
 50 55 60
 Asn Leu Arg Asp Leu Gln Ala Tyr Ala Leu Lys Gly Gly Leu Asn Gly
 65 70 75 80
 Ile Met Val Gly Gly Tyr Leu Thr Thr Gly Gly Arg Ser Pro Gln Asp
 85 90 95
 Asp Leu Gln Met Ile Gln Asp Leu Glu
 100 105

<210> 73
 <211> 384
 <212> DNA
 <213> Homo sapiens

<400> 73
 nntaccggca agatcctggc cgaagggtgac gtcgaggttt ctgaggctat cgactttgct
 60
 gcttggtatg tcgaccgagc cgaggagctc gagggcgctg acggtgcccc gtttgtgccg
 120
 ccacgagtga ccgtcgtcac cccgccgtgg aacttcgccc tgtctattac cgccggatcc
 180

acccttgccg ctctggccgc cggatcgta gtactactca agcccgctcc acaggcccgc
 240
 cactgtgctg ccgatcatctc tgaatgectg tgggaggctg ggatcccgcg ggacgttctg
 300
 cagctcgctg atgttgagga aaatgaggct ggtaaaccac tggtagacca ccccgaggtc
 360
 gatcgggtca tcttcacggg aggt
 384

<210> 74

<211> 128

<212> PRT

<213> Homo sapiens

<400> 74

Xaa	Thr	Gly	Lys	Ile	Leu	Ala	Glu	Gly	Asp	Val	Glu	Val	Ser	Glu	Ala
1			5					10						15	
Ile	Asp	Phe	Ala	Ala	Trp	Tyr	Val	Asp	Arg	Ala	Glu	Glu	Leu	Glu	Gly
	20						25						30		
Val	Asp	Gly	Ala	Gln	Phe	Val	Pro	Arg	Val	Thr	Val	Val	Thr	Pro	
	35					40				45					
Pro	Trp	Asn	Phe	Ala	Leu	Ser	Ile	Thr	Ala	Gly	Ser	Thr	Leu	Ala	Ala
	50				55					60					
Leu	Ala	Ala	Gly	Ser	Ser	Val	Leu	Leu	Lys	Pro	Ala	Pro	Gln	Ala	Arg
65				70					75					80	
His	Cys	Ala	Ala	Val	Ile	Ser	Glu	Cys	Leu	Trp	Glu	Ala	Gly	Ile	Pro
			85				90						95		
Arg	Asp	Val	Leu	Gln	Leu	Val	Asp	Val	Glu	Glu	Asn	Glu	Ala	Gly	Lys
		100					105					110			
His	Leu	Val	Ser	His	Pro	Glu	Val	Asp	Arg	Val	Ile	Leu	Thr	Gly	Gly
	115					120						125			

<210> 75

<211> 405

<212> DNA

<213> Homo sapiens

<400> 75

gaattcgtct cggaatacac gctggaaaat tcggccgaga tgtccgggggt gcgctcanac
 60
 cgcattgagg cgctggccga gctctatgcc gatcccaaga ccagggtggt gagcttctgg
 120
 accatgggct tcaaccagca caccgcgggc gtctggtgca acaatctcgt ctacaacatc
 180
 cacctgctga ccggaaaaat ctcgacgccc ggcaacagcc cgcttctcgt gaccgggcag
 240
 ccacggcct gggcacggc gcgcgaggtc ggtaccttct cgcacgcct gcccgcgac
 300
 atggtggtca ccagcaaggc gcaccgcgac atcgccgaga agatctggca gctgccggaa
 360
 ggaccagtcc ccgacaagcc cggctaccac gccgtgctgc agagc
 405

<210> 76

<211> 135
 <212> PRT
 <213> Homo sapiens

<400> 76
 Glu Phe Val Ser Glu Tyr Thr Leu Glu Asn Ser Ala Glu Met Ser Gly
 1 5 10 15
 Val Arg Ser Xaa Arg Ile Glu Ala Leu Ala Glu Leu Tyr Ala Asp Pro
 20 25 30
 Lys Thr Arg Val Val Ser Phe Trp Thr Met Gly Phe Asn Gln His Thr
 35 40 45
 Arg Gly Val Trp Cys Asn Asn Leu Val Tyr Asn Ile His Leu Leu Thr
 50 55 60
 Gly Lys Ile Ser Thr Pro Gly Asn Ser Pro Phe Ser Leu Thr Gly Gln
 65 70 75 80
 Pro Ser Ala Cys Gly Thr Ala Arg Glu Val Gly Thr Phe Ser His Arg
 85 90 95
 Leu Pro Ala Asp Met Val Val Thr Ser Lys Ala His Arg Asp Ile Ala
 100 105 110
 Glu Lys Ile Trp Gln Leu Pro Glu Gly Pro Val Pro Asp Lys Pro Gly
 115 120 125
 Tyr His Ala Val Leu Gln Ser
 130 135

<210> 77
 <211> 5816
 <212> DNA
 <213> Homo sapiens

<400> 77
 gagcggcgcc ctgctctggc cggttgetccc gctcctgctc ctgctgctgt cggcgcgggga
 60
 cggcgtgctgc gccgcgcagc ctcaggcccc ggggttacttg attgcagctc cctctgtttt
 120
 tcgcgcgggc gtggaggaag tcatcagcgt gaccatcttt aactctccaa gggaagtcac
 180
 ggtccaggct cagctggtgg cccagggtga gccggtggtg cagagccagg gagccatcct
 240
 ggataaagg acaatcaaac tcaagggtgcc cacgggcctc cggggccaag cgcttctgaa
 300
 agtgtggggc cgcggctggc aggcggagga gggggccctc tttcacaacc agacctcggt
 360
 gaccgtggac ggccggggcg cttctgtatt catccagacg gacaagcctg tgtacagacc
 420
 ccagcaccga gtgctcataa gcatcttcac cgtctctcca aatctgaggc ctgtcaacga
 480
 gaagctggaa gcctacatcc tggacccccg aggtctctcg atgatagagt ggagacactt
 540
 aaagccgttc tgetgcggca tcaccaacat gagcttcccc ttgtccgacc agcctgtgtt
 600
 gggagaatgg ttcatttttg ttgaaatgca aggccacgcg tacaacaagt cttttgaagt
 660
 tcagaagtat gtgttgccca agtttgagct tctgattgac ccgccccggt atatccaaga
 720

cctggacgcc tgtgagacag gcactgtgcy ggccaggtat acctttggga aacctgtggc
780
tgggtgcctta atgatcaaca tgactgttaa tgggttaggg tactacagcc acgaggtggg
840
acgccctgtc ctcagaacaa ccaagatcct cggctcccgg gacttcgaca tctgcgtgag
900
ggacatgatc ccagcggacg tccctgagca cttccggggc agggtcagca tctgggccat
960
ggtgaccagt gtggacggga gccagcaggt cgcgttcgat gactccaccc ccgtgcagag
1020
gcagctggtg gacatccggt actccaagga cagaggaag cagttcaagc cgggcctggc
1080
ctacgtgggg aaggtggagc tatectaccc cgatggcagc ccagctgagg ggggtgacggt
1140
ccagattaag gcagagctga caccaaagga taacatctac accagtgaag ttgtgtccca
1200
gcgtggacta gtgggggttg aaatcccctc catccccacg tcagcccagc acgtgtggct
1260
ggagaccaag gtgatggcac tgaacgggaa gcccgtgggg gctcagtacc tgcccagcta
1320
cctctccctc ggcagctggt actccccag ccagtgtac ctgcagctgc agccaccctc
1380
ccaccactg caggttgggg aagaagccta tttttctgtg aagtccacat gtccttgcaa
1440
ctttaccctg tactacgagg tggctgcacg gggcaatatt gtgctatcgg gccagcagcc
1500
tgcccacacc acccagcagc gaagcaagcg ggcggcccct gccctggaga aaccgattcg
1560
tttaacacac ctttctgaga cagagccccc accagcccca gaagctgagg tcgacgtgtg
1620
tgtgacctct cttcatctgg ccgtgacccc cagcatggtc ccccttggtc gcctgctggg
1680
cttctacgtc agggagaatg gagaaggggt cgccgacagc cttcagtttg cagtcgagac
1740
cttcttcgaa aaccaggttt cagtgcgta ttcagcaa at gagaccaac ctggggagg
1800
tgtgcacctg cggatcaggg ctgcaagggg cagctgtgtg tgcgtcgccg cagttgataa
1860
gagtgtctac ctgctcaggt ctgggttccg gctgactcct gccaggttt tccaggaact
1920
ggaagattat gatgtttctg attcctttgg cgtgtccagg gaggatggtc ctttttggtg
1980
ggctgggctg acggcacaac gacgccggcg ctctctgtc tcccgtggc cttggggcat
2040
caccaaggac tctgggtttg ctttcaccga aacgggactg gtggtgatga ccgaccgagt
2100
gagcctgaac caccggcagg acggtggcct ctacaccgat gaggtgtcc ccgctttcca
2160
gccccacaca gggagcctgg tggcagtggc tccttcagg ccccccca gaacagagaa
2220
gagaaaaagg actttcttcc ccgaaacatg gatttggcat tgtctcaaca tcagtgaccc
2280
atctggtgag gggacactca gtgtgaaggt cccggactcc atcaccagct ggggtgggtga
2340

ggccgtggcc ctgtccacct ctcagggett aggcacgcc gagccctccc tgctgaagac
2400
cttcaagccc ttcttcgtgg acttcatgct ccccgtcttc atcatccgtg gggagcaggt
2460
caagatcccc ctcagtgtct acaactacat gggcacctgc gctgaggtgt acatgaagct
2520
ctcggttccc aagggcattc agtttggttg gcatcctggc aaacgccatg tgaccaagaa
2580
gatgtgtgtg gccccggggg aggctgagcc catctgggtc gttctgtcct tcagcgacct
2640
gggactcaac aacatcacgg ccaaagccct tgcttacgga gacacaaatt gctgccggga
2700
tgggaggtcc agcaaacacc ctgaggagaa tcacgccgac aggaggggtcc ccatcggggg
2760
ggatcacgtc aggcgcagtg tgatggttga ggcggaagga gtccccggg cgtacaccta
2820
cagcgcattc ttctgtccca gtgagagagt ccacatctcc accccaaca agtatgagtt
2880
ccagtatgtg cagcggccac tgcgcctcac ccgctttgat gtggctgtgc gagtcacaa
2940
tgatgcccgt gtggccttgt cttctggggc ccaggacaca gcaggcatga tcgagatcgt
3000
cctggggggg catcagaaca ccaggctcatg gatctccacc agcaagatgg gagagcccgt
3060
ggccagtgca cacacggcca agatcctctc ctgggatgaa ttcagaacat tctggatcag
3120
ctggcgtggt ggccttatcc aggttggcca tggccagag ccatccaatg agtctgtcat
3180
tgtggcctgg accctcccga ggccaccaga ggtccagttc attggctttt ccaccggctg
3240
gggtccatg ggtgaattcc gaatctggag gaagatggag gtggacgaga gctacagcga
3300
ggccttcacc tgggggtccc acacggcgcc atccctgggt ctgagcgagc caccgcctcc
3360
atcatcgggg acgtcatggg gccaaacctg aaccacctca acaacctcct ggggtgccg
3420
tttggtgtg gagagcagaa catgatccac tttgcacca acgtctttgt cttgaagtat
3480
cttcagaaaa cccagcagct cagccctgag gtggagagag agaccaccga ctacctagta
3540
caaggctacc agcgccagct gacctacaag cgccaggatg gtcctacag cgcgtttggg
3600
gagcgggacg catcggggag catgtggctc acagcctttg tcctgaagtc cttcgcacag
3660
gctcgagct ttatcttcgt ggacccccgg gagctgggtg ccgccaagag ctggatcatc
3720
cagcagcagc aggcgatgg ctccttctcg gccgtgggca gggctcctgaa caaggacatc
3780
caggggtggga tccacggcat tgtcccgtg acagcctacg tgggtggtgc tctcctggaa
3840
acaggcacag cctcagagga ggagagagc tccactgaca aagcgaggca cttcctggag
3900
tctgtgcgc ccctggccat ggaccttat agctgtgccc tgactaccta cgcgtgacc
3960

ctgctccgca gcccggcagc ccctgaggca ctgcgcaagc tccgtagcct ggccatcatg
4020
cgagatgggg tcacccactg gagcctgtca aattcctggg acgtggacaa gggcacattc
4080
ttgagcttca gtgacagggc ctctcagtca gtggtctcgg ccgaggtgga aatgacagcc
4140
tacgcccttc tgacctacac tctgctgggt gacgtggctg ccgccctgcc tgtggtgaag
4200
tggctgtccc agcagcgaag tgcacttggg ggcttctcct ccactcagga cacctgcgtg
4260
gctctgcagg ccttggtgga atatgccatc ttgtcctatg ctggaggcat caacctcact
4320
gtctccctgg cctccaccaa cctggactac caggaaacct tcgagctgca caggaccaac
4380
cagaagggtc tgacagacagc agcgatcccc agcctcccca cggggctggt tgtgagtgcc
4440
aagggggacg gctgctgctt gatgcagatt gatgtcacct acaatgtgcc tgaccgggtg
4500
gccaagccag ctttccagct gctcgtaagc ctccaggagc ctgaggccca gggacgcccg
4560
cccccatgc ctgcctccgc agctgagggc tcccaggagc actggccccc agctgacgat
4620
gatgacccag cggccgatca gcatcaccag gaatacaagg tgatgctgga ggtgtgcacc
4680
aggtggctgc atgcagggtc ttccaatatg gctgtcctgg aggtgcccct gctgtcaggc
4740
ttccgggcag acatcgagag cctggagcag ctgctccttg acaagcacat ggggatgaag
4800
aggatgaag tggctggacg ccgagtgtc ttctactttg atgagatccc cagccgggtc
4860
ctgacgtgcg tgcggttccg tgctctccgg gagtgcgtgg tgggcaggac gtcggcgctg
4920
ccagtctccg tgtacgacta ctacgaaccc gccttcgagg ccactcgctt ctacaacgtc
4980
agcacgcaca gccactcgc ccgggaactg tgcgccggac ccgctgcaa cgaagtggag
5040
cgcgcccctg ccggggccc gggctggtc ccggcgagt cgggccctgc cgtggcccct
5100
gaggaggggg cggcgatcgc gcgatgcggc tgcgaccacg actgcggcgc ccaggggaac
5160
ccggtgtgcg gctccgacgg ggtggtctac gccagcgctt gccgcctgcg ggaggccgcc
5220
tgccgccagg ccgcgcccct ggagcccgcg cctcccagct gctgcgccct cgagcagcg
5280
ctgcggcct cgtcgtcctc cacctacggg gatgacctg cttctgtggc cccggggcct
5340
ttacagcagg acgtgaagct gaatggagcc ggccttgagg tggaggactc agaccctgag
5400
cctgaagggg agggcgagga cagggtcaca gccgggcctc ggccctcctgt gagcagcggg
5460
aacctggaaa gcagcaccca gagcgccagc ccgttcaca gatggggcca gactccggcc
5520
cctcagagac atagtggccg ggtggtgggg gccacaggc cagggttct gagccctgtc
5580

ttctgtctaca gccacgcctt tcagagtggg ggggaggagg gtttatggat gtcaaacacc
 5640
 tgcaccttga gataatccta caaccacatg cagttgtggg accgcagttt ggtcctgggg
 5700
 accattcata cccacacacc cagcttgtgc ctgtgggttaa catctcagaa aactctggta
 5760
 aatgatcact ccaggatatt gacacgaata cacgttactg atcttactca catgtt
 5816

<210> 78

<211> 799

<212> PRT

<213> Homo sapiens

<400> 78

Ile	Pro	Asn	Leu	Glu	Glu	Asp	Gly	Gly	Gly	Arg	Glu	Leu	Gln	Arg	Gly	1	5	10	15
Leu	His	Leu	Gly	Val	Pro	His	Gly	Ala	Ile	Pro	Gly	Ser	Glu	Arg	Ala	20	25	30	
Thr	Ala	Ser	Ile	Ile	Gly	Asp	Val	Met	Gly	Pro	Thr	Leu	Asn	His	Leu	35	40	45	
Asn	Asn	Leu	Leu	Arg	Leu	Pro	Phe	Gly	Cys	Gly	Glu	Gln	Asn	Met	Ile	50	55	60	
His	Phe	Ala	Pro	Asn	Val	Phe	Val	Leu	Lys	Tyr	Leu	Gln	Lys	Thr	Gln	65	70	75	80
Gln	Leu	Ser	Pro	Glu	Val	Glu	Arg	Glu	Thr	Thr	Asp	Tyr	Leu	Val	Gln	85	90	95	
Gly	Tyr	Gln	Arg	Gln	Leu	Thr	Tyr	Lys	Arg	Gln	Asp	Gly	Ser	Tyr	Ser	100	105	110	
Ala	Phe	Gly	Glu	Arg	Asp	Ala	Ser	Gly	Ser	Met	Trp	Leu	Thr	Ala	Phe	115	120	125	
Val	Leu	Lys	Ser	Phe	Ala	Gln	Ala	Arg	Ser	Phe	Ile	Phe	Val	Asp	Pro	130	135	140	
Arg	Glu	Leu	Ala	Ala	Ala	Lys	Ser	Trp	Ile	Ile	Gln	Gln	Gln	Gln	Ala	145	150	155	160
Asp	Gly	Ser	Phe	Leu	Ala	Val	Gly	Arg	Val	Leu	Asn	Lys	Asp	Ile	Gln	165	170	175	
Gly	Gly	Ile	His	Gly	Ile	Val	Pro	Leu	Thr	Ala	Tyr	Val	Val	Val	Ala	180	185	190	
Leu	Leu	Glu	Thr	Gly	Thr	Ala	Ser	Glu	Glu	Glu	Arg	Gly	Ser	Thr	Asp	195	200	205	
Lys	Ala	Arg	His	Phe	Leu	Glu	Ser	Ala	Ala	Pro	Leu	Ala	Met	Asp	Pro	210	215	220	
Tyr	Ser	Cys	Ala	Leu	Thr	Thr	Tyr	Ala	Leu	Thr	Leu	Leu	Arg	Ser	Pro	225	230	235	240
Ala	Ala	Pro	Glu	Ala	Leu	Arg	Lys	Leu	Arg	Ser	Leu	Ala	Ile	Met	Arg	245	250	255	
Asp	Gly	Val	Thr	His	Trp	Ser	Leu	Ser	Asn	Ser	Trp	Asp	Val	Asp	Lys	260	265	270	
Gly	Thr	Phe	Leu	Ser	Phe	Ser	Asp	Arg	Val	Ser	Gln	Ser	Val	Val	Ser	275	280	285	
Ala	Glu	Val	Glu	Met	Thr	Ala	Tyr	Ala	Leu	Leu	Thr	Tyr	Thr	Leu	Leu	290	295	300	
Gly	Asp	Val	Ala	Ala	Ala	Leu	Pro	Val	Val	Lys	Trp	Leu	Ser	Gln	Gln				

305						310										320
Arg	Asn	Ala	Leu	Gly	Gly	Phe	Ser	Ser	Thr	Gln	Asp	Thr	Cys	Val	Ala	
				325					330						335	
Leu	Gln	Ala	Leu	Ala	Glu	Tyr	Ala	Ile	Leu	Ser	Tyr	Ala	Gly	Gly	Ile	
			340					345					350			
Asn	Leu	Thr	Val	Ser	Leu	Ala	Ser	Thr	Asn	Leu	Asp	Tyr	Gln	Glu	Thr	
		355					360					365				
Phe	Glu	Leu	His	Arg	Thr	Asn	Gln	Lys	Val	Leu	Gln	Thr	Ala	Ala	Ile	
	370					375					380					
Pro	Ser	Leu	Pro	Thr	Gly	Leu	Phe	Val	Ser	Ala	Lys	Gly	Asp	Gly	Cys	
385					390					395					400	
Cys	Leu	Met	Gln	Ile	Asp	Val	Thr	Tyr	Asn	Val	Pro	Asp	Pro	Val	Ala	
			405						410					415		
Lys	Pro	Ala	Phe	Gln	Leu	Leu	Val	Ser	Leu	Gln	Glu	Pro	Glu	Ala	Gln	
		420						425					430			
Gly	Arg	Pro	Pro	Pro	Met	Pro	Ala	Ser	Ala	Ala	Glu	Gly	Ser	Arg	Gly	
	435						440					445				
Asp	Trp	Pro	Pro	Ala	Asp	Asp	Asp	Asp	Pro	Ala	Ala	Asp	Gln	His	His	
	450					455					460					
Gln	Glu	Tyr	Lys	Val	Met	Leu	Glu	Val	Cys	Thr	Arg	Trp	Leu	His	Ala	
465					470					475					480	
Gly	Ser	Ser	Asn	Met	Ala	Val	Leu	Glu	Val	Pro	Leu	Leu	Ser	Gly	Phe	
			485						490					495		
Arg	Ala	Asp	Ile	Glu	Ser	Leu	Glu	Gln	Leu	Leu	Leu	Asp	Lys	His	Met	
		500						505					510			
Gly	Met	Lys	Arg	Tyr	Glu	Val	Ala	Gly	Arg	Arg	Val	Leu	Phe	Tyr	Phe	
	515						520					525				
Asp	Glu	Ile	Pro	Ser	Arg	Cys	Leu	Thr	Cys	Val	Arg	Phe	Arg	Ala	Leu	
	530					535					540					
Arg	Glu	Cys	Val	Val	Gly	Arg	Thr	Ser	Ala	Leu	Pro	Val	Ser	Val	Tyr	
545					550					555					560	
Asp	Tyr	Tyr	Glu	Pro	Ala	Phe	Glu	Ala	Thr	Arg	Phe	Tyr	Asn	Val	Ser	
			565						570					575		
Thr	His	Ser	Pro	Leu	Ala	Arg	Glu	Leu	Cys	Ala	Gly	Pro	Ala	Cys	Asn	
			580					585					590			
Glu	Val	Glu	Arg	Ala	Pro	Ala	Arg	Gly	Pro	Gly	Trp	Phe	Pro	Gly	Glu	
	595						600					605				
Ser	Gly	Pro	Ala	Val	Ala	Pro	Glu	Glu	Gly	Ala	Ala	Ile	Ala	Arg	Cys	
	610					615					620					
Gly	Cys	Asp	His	Asp	Cys	Gly	Ala	Gln	Gly	Asn	Pro	Val	Cys	Gly	Ser	
625					630					635					640	
Asp	Gly	Val	Val	Tyr	Ala	Ser	Ala	Cys	Arg	Leu	Arg	Glu	Ala	Ala	Cys	
			645													

```

          740          745          750
Thr Pro Ala Pro Gln Arg His Ser Gly Arg Val Val Gly Ala His Arg
          755          760          765
Pro Gly Leu Leu Ser Pro Val Phe Val Tyr Ser Pro Ala Phe Gln Ser
          770          775          780
Gly Gly Glu Glu Gly Leu Trp Met Ser Asn Thr Cys Thr Leu Arg
785          790          795

```

<210> 79
 <211> 346
 <212> DNA
 <213> Homo sapiens

```

<400> 79
acgcgttgcc gtcggggaag tggtagatgt aaggctcttc catttctgcg gcggcttctt
60
cgctgatagg cgcgaggttc atgccgccaa agccccgggaa acccacgctg taggcacccg
120
ccgatgcaac ggcatacaata ctcaaggcca tgaagtcggtt ggtgcggtc tggaagtact
180
tcaccgcgcc ttccgacaag cccacgtcct tgagcaggaa gtcgcggtag ctggtggccg
240
ccagatactc ggctttttct tcggcggact tgcccggcag gtaatccttg ggcgcgacgt
300
gcatggcgat caatgcccg ctgtccgctt ccggnccnnnc cnnccnn
346

```

<210> 80
 <211> 101
 <212> PRT
 <213> Homo sapiens

```

<400> 80
Met His Val Ala Pro Lys Asp Tyr Leu Pro Gly Lys Ser Ala Glu Glu
1          5          10          15
Lys Ala Glu Tyr Leu Ala Ala Thr Ser Tyr Arg Asp Phe Leu Leu Lys
20          25          30
Asp Val Gly Leu Ser Glu Gly Ala Val Lys Tyr Phe Gln Ser Arg Thr
35          40          45
Asn Asp Phe Met Ala Leu Ser Ile Asp Ala Val Ala Ser Ala Asp Ala
50          55          60
Tyr Ser Val Gly Phe Pro Gly Phe Gly Gly Met Asn Leu Ala Pro Ile
65          70          75          80
Ser Glu Glu Ala Ala Ala Glu Met Glu Glu Pro Tyr Ile Tyr His Phe
85          90          95
Pro Asp Gly Asn Ala
100

```

<210> 81
 <211> 429
 <212> DNA
 <213> Homo sapiens

<400> 81

nngcttttct cncctcgca cctgatctgc tgcctgacgtg cgggaggggc attaaaagat
 60
 tgacttaaat ttgataccag aaatgtcagt gtggctttgc ttgaacagcc gcgtgcgggg
 120
 cctctatgag tggaatccag tctcatggcc ccccccattg ctctgtttac cctggaggag
 180
 gctactccat gaggttggt cgggtggcac tgcctcgggg ctgcatgtac atgtgtgtgc
 240
 atgnntgtgt gcatgtgcgt gtgcacgtgt nngtgtgtgc ccgtgtgcat gtgcccgtgt
 300
 gcgtgtctgt gctgtgtgtg tgcgtgcatg tgtgcgtgtc tgtgcgtgnc tgtgtgtgtg
 360
 cgtgcatgtg tgtgcacgtg tgtgcatgtg cacgtgtgcg caccctgtgt catgtccgtg
 420
 tgcacgcgt
 429

<210> 82

<211> 79

<212> PRT

<213> Homo sapiens

<400> 82

Gly	Trp	Leu	Arg	Trp	His	Cys	Leu	Gly	Ala	Ala	Cys	Thr	Cys	Val	Cys
1				5				10					15		
Met	Xaa	Val	Cys	Met	Cys	Val	Cys	Thr	Cys	Xaa	Cys	Val	Pro	Val	Cys
			20				25					30			
Met	Cys	Pro	Cys	Ala	Cys	Leu	Cys	Cys	Val	Cys	Ala	Cys	Met	Cys	Ala
			35				40					45			
Cys	Leu	Cys	Val	Xaa	Val	Cys	Val	Arg	Ala	Cys	Val	Cys	Thr	Cys	Val
	50					55					60				
His	Val	His	Val	Cys	Ala	Pro	Val	Cys	Met	Ser	Val	Cys	Thr	Arg	
65					70					75					

<210> 83

<211> 411

<212> DNA

<213> Homo sapiens

<400> 83

gtgcacttct tcggacgagg gtccttcggc gcccgcgata acgagagggg gcggagattt
 60
 cgtgcgctct ttaaaaagcg caaaggcttt tataagttcg atgtgctttt tttcaggccc
 120
 ggaaagacgc gatccgtaga taaaataggg ccgctttatc gcaaattggc ttatgggttac
 180
 cgtttccgca tcgatttgca gctgcggaaa aaaaggcctg tggtcgatgc cgttgtggat
 240
 gatttcgatt ttcgctttgt cgatgccaaag ccctatcaaa tcgtttcgca aaaactcgct
 300
 cgcggcgacg atcttttgca cctttgtaag ccccgcttt atctgcatgc gcatcatgac
 360
 gtcgctcttt tctttgagca tgacggaaac gatggaattg acgacggcga c
 411

<210> 84
 <211> 127
 <212> PRT
 <213> Homo sapiens

<400> 84
 Met Leu Lys Glu Lys Ser Asp Arg Met Met Arg Met Gln Ile Lys Arg
 1 5 10 15
 Gly Leu Thr Lys Val Gln Lys Ile Val Ala Ala Ser Glu Phe Leu Arg
 20 25 30
 Asn Asp Leu Ile Gly Leu Gly Ile Asp Lys Ala Lys Ile Glu Ile Ile
 35 40 45
 His Asn Gly Ile Asp His Arg Pro Phe Phe Pro Gln Leu Gln Ile Asp
 50 55 60
 Ala Glu Thr Val Thr Ile Lys Pro Phe Ala Ile Lys Arg Pro Tyr Phe
 65 70 75 80
 Ile Tyr Gly Ser Arg Leu Ser Gly Pro Glu Lys Lys His Ile Glu Leu
 85 90 95
 Ile Lys Ala Phe Ala Leu Phe Lys Glu Arg Thr Lys Ser Pro His Pro
 100 105 110
 Leu Val Ile Ala Gly Ala Glu Gly Pro Ser Ser Glu Glu Val His
 115 120 125

<210> 85
 <211> 333
 <212> DNA
 <213> Homo sapiens

<400> 85
 nnacgcgtgc cgcgcaaggg aaccatgttc gtgtcggtgc gcgaaaccga caaggcgcgc
 60
 atcctcgacg cggtgaaact gctgagttcg ctcggttca aggtgatcgc cacctcgggc
 120
 acccagcgtt tcctggtgga gaacggagta ccggcggaag agatcaacaa ggtgctggaa
 180
 ggccgcccgc acattgttga cgcaattacc aacggcgagg tgcaactcgt tttcaatacc
 240
 accgaggggc cacaggcgtt ggctgacagc cgctcggttc gacgcgctgc cctcttgcat
 300
 aaagtgccat attacaccac tctttcaggt gca
 333

<210> 86
 <211> 111
 <212> PRT
 <213> Homo sapiens

<400> 86
 Xaa Arg Val Pro Arg Lys Gly Thr Met Phe Val Ser Val Arg Glu Thr
 1 5 10 15
 Asp Lys Ala Arg Ile Leu Asp Ala Val Lys Leu Leu Ser Ser Leu Gly
 20 25 30
 Phe Lys Val Ile Ala Thr Ser Gly Thr Gln Arg Phe Leu Val Glu Asn

ATTORNEY DOCKET NO.: 15966-543

```

      35          40          45
Gly Val Pro Ala Glu Lys Ile Asn Lys Val Leu Glu Gly Arg Pro His
  50          55          60
Ile Val Asp Ala Ile Thr Asn Gly Glu Val Gln Leu Val Phe Asn Thr
  65          70          75          80
Thr Glu Gly Pro Gln Ala Leu Ala Asp Ser Arg Ser Leu Arg Arg Ala
      85          90          95
Ala Leu Leu His Lys Val Pro Tyr Tyr Thr Thr Leu Ser Gly Ala
      100          105          110

```

<210> 87

<211> 355

<212> DNA

<213> Homo sapiens

<400> 87

```

acgcgtgagg aaatgggggc cgcaggcctg gatcgcaagg tatggcagtg cccggtcgtc
60
ctcctgagcg atgttcattc ggtaggggta caggggtgacg ggcgtactta tggttctccc
120
attgtgcttc gcccggtgac gagtgaggac gccatgactg cggactgggc acgtatccca
180
tatgacgtac tggaaaagat ctcgactcgc attacgaatg cgtgtccgca aatcaaccgg
240
gtggtactcg atatcacatc taaaccgccg gccaccatcg agtgggaatg agccccgtct
300
caccgtgaac atgacatggc ccgcaccttt cttggggcgg gccatgccgt gttag
355

```

<210> 88

<211> 96

<212> PRT

<213> Homo sapiens

<400> 88

```

Thr Arg Glu Glu Met Gly Ala Ala Gly Leu Asp Arg Lys Val Trp Gln
  1          5          10          15
Cys Pro Val Val Leu Leu Ser Asp Val His Ser Val Gly Val Gln Gly
      20          25          30
Asp Gly Arg Thr Tyr Gly Ser Pro Ile Val Leu Arg Pro Val Thr Ser
      35          40          45
Glu Asp Ala Met Thr Ala Asp Trp Ala Arg Ile Pro Tyr Asp Val Leu
      50          55          60
Glu Lys Ile Ser Thr Arg Ile Thr Asn Ala Cys Pro Gln Ile Asn Arg
      65          70          75          80
Val Val Leu Asp Ile Thr Ser Lys Pro Pro Ala Thr Ile Glu Trp Glu
      85          90          95

```

<210> 89

<211> 351

<212> DNA

<213> Homo sapiens

<400> 89

ATTORNEY DOCKET NO.: 15966-543

nacgcgtcaa caccaggcta cgggtgggtat gatcatgata agggctggga cccgcaggag
 60
 agggcaagag ttatataata gacgggtgta tgggtttaag gtatacatgg tgatgagaca
 120
 ggtgattgga taaggccaga ggggtgggag ggggtctgcc cctgctgaag cctgggtggg
 180
 cccaggtctg tgatctggga ccggaacaac acatctgctc tgggcctgct ggatgtggcg
 240
 caagccctgg aacagaacca cagcctcaag tccatgccgc tgccactgaa tgacgtaacc
 300
 caggctcatc gcagccggcc agaactcaca actcgagcgg tccatcagat c
 351

<210> 90
 <211> 61
 <212> PRT
 <213> Homo sapiens

<400> 90
 Ser Leu Val Gly Pro Arg Ser Val Ile Trp Asp Arg Asn Asn Thr Ser
 1 5 10 15
 Ala Leu Gly Leu Leu Asp Val Ala Gln Ala Leu Glu Gln Asn His Ser
 20 25 30
 Leu Lys Ser Met Pro Leu Pro Leu Asn Asp Val Thr Gln Ala His Arg
 35 40 45
 Ser Arg Pro Glu Leu Thr Thr Arg Ala Val His Gln Ile
 50 55 60

<210> 91
 <211> 327
 <212> DNA
 <213> Homo sapiens

<400> 91
 nggtgggtcc tctcgcagga caggtaatcc tgagacctac gcagctccct ggagctctgc
 60
 acgtcgaagc ccagcagggc ctccctgcagg tccctggggc agccagcaca cacaaagtcc
 120
 cggaaggggc tgtagactcc ctgccagcgg ctttcccggg gaaggcacc acgcccagct
 180
 gcctcttgca ggtactgctc ggggtctggtg ggagggcagc cgtgtccagc acaccctgtg
 240
 tgtgcagtcc tctccctgcc ccaactgccga acgagccctc cacggtgaag ccattgggga
 300
 acgtgacctt gcccttcccc atgaggt
 327

<210> 92
 <211> 107
 <212> PRT
 <213> Homo sapiens

<400> 92
 Met Gly Lys Gly Lys Val Thr Phe Pro Asn Gly Phe Thr Val Glu Gly

ATTORNEY DOCKET NO.: 15966-543

```

      1           5           10           15
Ser Phe Gly Ser Gly Ala Gly Arg Gly Leu His Thr Gln Gly Val Leu
      20           25           30
Asp Thr Ala Ala Leu Pro Pro Asp Pro Ser Ser Thr Cys Lys Arg Gln
      35           40           45
Leu Gly Val Gly Ala Phe Pro Gly Lys Ala Ala Gly Arg Glu Ser Thr
      50           55           60
Ala Pro Ser Gly Thr Leu Cys Val Leu Ala Ala Pro Gly Thr Cys Arg
      65           70           75           80
Arg Pro Cys Trp Ala Ser Thr Cys Arg Ala Pro Gly Ser Cys Val Gly
      85           90           95
Leu Arg Ile Thr Cys Pro Ala Arg Gly Pro Thr
      100          105

```

<210> 93
 <211> 394
 <212> DNA
 <213> Homo sapiens

```

<400> 93
nccgcgtacg acaagcagta cctcgagggg cggtacggtg cggacccata cctgagcaac
60
atgctcgaat gggacggcgg acatgagcaa taggccgcca aagcacagcg agaggaaggc
120
ggcgcgctct gcggcctcgc ggatgaagga ggggtggtcg cgcaagagac gccccgcgcc
180
ttttgtaccg ataaatcccc ggcgccacct gatcgtgacc gaggggtacga aaacggaacc
240
gctctatttc gaggtatca ggttgctgt caacaaccgt taccacggcc agtgggtgac
300
aatggaagtt gtcgttaccg gcaagcatac caggggactt ctcgatcgtg cagtcactct
360
ggcggaagaa agtgccacag gattcactca cgta
394

```

<210> 94
 <211> 109
 <212> PRT
 <213> Homo sapiens

```

<400> 94
Met Leu Ala Gly Asn Asp Asn Phe His Cys His Pro Leu Ala Val Ile
1           5           10           15
Thr Val Val Asp Thr Gln Pro Asp Ser Leu Glu Ile Glu Arg Phe Arg
      20           25           30
Phe Arg Thr Leu Gly His Asp Gln Val Ala Pro Gly Ile Tyr Arg Tyr
      35           40           45
Lys Arg Arg Gly Ala Ser Leu Ala Arg Pro Pro Leu Leu His Pro Arg
      50           55           60
Gly Arg Arg Ala Arg Arg Leu Pro Leu Ala Val Leu Trp Arg Pro Ile
      65           70           75           80
Ala His Val Arg Arg Pro Ile Arg Ala Cys Cys Ser Gly Met Gly Pro
      85           90           95
His Arg Asn Ala Pro Arg Gly Thr Ala Cys Arg Thr Arg

```


100

105

<210> 95
 <211> 531
 <212> DNA
 <213> Homo sapiens

<400> 95
 ggtacctctt ccaagtacct tctaaatgaa acactcaaga gagggtact caggaaactt
 60
 tgcttggtatc ctaaaatgga ctgggtcttg gtgtgtaacc ccggtgaagt tatagcctcc
 120
 ccaaattgag gtgacagaag gaagacaaga ggtgtaagct ggagagggaa gggaagaaat
 180
 cagtggcttt ggccagcctc tgtgccaccc agtacgacag aggagtggga actggccctc
 240
 tggggctctg cttggccata ggcactgcac attgtgccac ctgctcatca cctcctctag
 300
 tctcacactg agcatcggag tacctgttgt gcagacagga aaactgagga gctctgagag
 360
 gctgagcatg gagctcacc catgccatag ggtgtgggaa gagggcacag gaggcctcat
 420
 ccatggggga aagggttgag gatggacatg ggtggggaga gggcatagac atcccttcc
 480
 aatctctgtt cccaccacat ttcataggag atgagttagg agatgacagc t
 531

<210> 96
 <211> 124
 <212> PRT
 <213> Homo sapiens

<400> 96
 Met Arg Pro Pro Val Pro Ser Ser His Thr Leu Trp His Gly Val Ser
 1 5 10 15
 Ser Met Leu Ser Leu Ser Glu Leu Leu Ser Phe Pro Val Cys Thr Thr
 20 25 30
 Gly Thr Pro Met Leu Ser Val Arg Leu Glu Glu Val Met Ser Arg Trp
 35 40 45
 His Asn Val Gln Cys Leu Trp Pro Ser Arg Ala Pro Glu Gly Gln Phe
 50 55 60
 Pro Leu Leu Cys Arg Thr Gly Trp His Arg Gly Trp Pro Lys Pro Leu
 65 70 75 80
 Ile Ser Ser Leu Pro Ser Pro Ala Tyr Thr Ser Cys Leu Pro Ser Val
 85 90 95
 Thr Ser Ile Trp Gly Gly Tyr Asn Phe Thr Gly Val Thr His Pro Arg
 100 105 110
 Pro Val His Phe Arg Ile Gln Ala Lys Phe Pro Glu
 115 120

<210> 97
 <211> 405
 <212> DNA
 <213> Homo sapiens

<400> 97

ggctcgggcca gtcgaacttc attcccgcctt cgaggggtctt gctgcggatg ggccgtacgc
 60
 tcgcggtgcc ttgcgcgcgg gctggttaggt ggagaagccg cgcgagtacg cgccgtagag
 120
 cgacatcgtg tctgagacgt cgaagctcag gccagctttt ggctccagg cgcgctcgg
 180
 cggtcgcccc tcttgcgga attgattcag cgcaatcccc gccatcacat gccagcgctt
 240
 gtccagggtc atgaaatcct gggcataggc gcgcgaggag cgcagcggcg aattggacag
 300
 gcgctcgata ttgggcgtga tgtccgaaga cgggaacggg acccgggggg agaagacgtt
 360
 gcccgggaaa agatcccccg acgccatcgt ggtgtcgacc gagat
 405

<210> 98

<211> 122

<212> PRT

<213> Homo sapiens

<400> 98

Met	Ala	Ser	Gly	Asp	Leu	Phe	Pro	Gly	Asn	Val	Phe	Ser	Pro	Arg	Val
1			5					10					15		
Pro	Phe	Pro	Ser	Ser	Asp	Ile	Thr	Pro	Asn	Ile	Glu	Arg	Leu	Ser	Asn
		20					25					30			
Ser	Pro	Leu	Arg	Ser	Ser	Arg	Ala	Tyr	Ala	Gln	Asp	Phe	Met	Thr	Leu
	35					40				45					
Asp	Lys	Arg	Trp	His	Val	Met	Ala	Gly	Ile	Ala	Leu	Asn	Gln	Leu	Pro
50					55					60					
Gln	Glu	Gly	Gly	Pro	Thr	Glu	Arg	Ala	Trp	Thr	Pro	Lys	Leu	Gly	Leu
65				70				75					80		
Ser	Phe	Asp	Val	Ser	Asp	Thr	Met	Ser	Leu	Tyr	Gly	Ala	Tyr	Ser	Arg
			85				90					95			
Gly	Phe	Ser	Thr	Tyr	Gln	Pro	Ala	Arg	Lys	Ala	Pro	Arg	Ala	Tyr	Gly
		100				105					110				
Pro	Ser	Ala	Ala	Arg	Pro	Ser	Lys	Arg	Glu						
		115				120									

<210> 99

<211> 545

<212> DNA

<213> Homo sapiens

<400> 99

acgcgtccgc tcccgatgct gttgcgagc tgcgctcagc gatgacggta ctcacccatc
 60
 tgccccgacg acccagcaaa cgccccggc tgttcctcat tgaccacgcc gaccggatcg
 120
 tcgatcccat cactcgggat ttgctggaat ccctgggtcg cgaagccggc gagcgtcggg
 180
 tgatcttggg tgcccagcgt cgcggctgca tcgattgggt cccccacag atcatccaca
 240

acctggccga acaccatttt gagtcgtcct ctggaggtac tcgatgatga ctgaacgttc
 300
 ccattccacg atcagggttaa ggtggccggc ggtggtggtt ctgcctctcg ttccgctgct
 360
 ggtggtcgcc ggattggtcc gggacgacct ggcataccac cgaccgggtg ggccgggtga
 420
 aagcggccgt cgtcaacgag gacaaggccg tcaaggtgcg tggacaactg gttccgatgg
 480
 gccgccaact caccgccgcc ttgatggact ctggctcgca caccactgat ggccacaccg
 540
 tcgac
 545

<210> 100

<211> 101

<212> PRT

<213> Homo sapiens

<400> 100

Met	Gly	Thr	Phe	Ser	His	His	Arg	Val	Pro	Pro	Glu	Asp	Asp	Ser	Lys
1				5					10					15	
Trp	Cys	Ser	Ala	Arg	Leu	Trp	Met	Ile	Cys	Gly	Glu	Ser	Gln	Ser	Met
			20					25					30		
Arg	Pro	Arg	Arg	Trp	Ala	Pro	Lys	Ile	Thr	Ala	Ala	Ser	Pro	Ala	Ser
			35				40					45			
Arg	Thr	Arg	Asp	Ser	Ser	Lys	Ser	Arg	Val	Met	Gly	Ser	Thr	Ile	Arg
	50					55					60				
Ser	Ala	Trp	Ser	Met	Arg	Asn	Ser	Arg	Gly	Arg	Leu	Leu	Gly	Arg	Arg
65					70				75					80	
Gly	Arg	Trp	Val	Ser	Thr	Val	Ile	Ala	Glu	Arg	Ser	Ser	Ser	Thr	Thr
			85						90					95	
Ser	Gly	Ala	Asp	Ala											
															100

<210> 101

<211> 619

<212> DNA

<213> Homo sapiens

<400> 101

ngcgcgccac agaagtgaag aagtacggtc tcaagcccgt gggagcgttc ttctgtccg
 60
 tcatgcacga tccggacttc gatccgatcc ccatgggtgaa caaggagctt gacgccttcg
 120
 aagctgccgg ggggtgactat ctcatcctcg ccacggattc cggacgcaag ggatacacga
 180
 ccgcccgtcc tcacgaggcc ggcggaaaac gctattacca acctggacca gatccgcgaa
 240
 gtctgcgcca gccgcaacgt caccgcctgt ctacaccccc attggggaac gatgggtccag
 300
 aaccgtgacg aagtgatccg cgtgctcgag aactcctega tcgggctgtg cctggacact
 360
 ggtcatctgg cctgtggtgg taccgatgtc gttgagctgg tgcgtaagta cgccaaccgc
 420

gtcgacattg tccacgcca agatgtccat aaggagatgg ccgacaagct tttgctggc
 480
 gagatcacct ggtccgaagg cattcgcc gccgatgttcg caccatcgg cgacgggtgat
 540
 atcgactttg cagccatcgt gaggtcctt gatgaagccg gggtcgatgg ttattacgtc
 600
 ctagagcagg acatcatga
 619

<210> 102

<211> 173

<212> PRT

<213> Homo sapiens

<400> 102

Thr	Arg	Ser	Leu	Thr	Pro	Ser	Lys	Leu	Pro	Gly	Val	Thr	Ile	Ser	Ser
1				5					10					15	
Ser	Pro	Arg	Ile	Pro	Asp	Ala	Arg	Asp	Thr	Arg	Pro	Pro	Val	Leu	Thr
			20					25					30		
Arg	Pro	Ala	Glu	Asn	Ala	Ile	Thr	Asn	Leu	Asp	Gln	Ile	Arg	Glu	Val
		35					40				45				
Cys	Ala	Ser	Arg	Asn	Val	Thr	Ala	Cys	Leu	His	Pro	His	Trp	Gly	Thr
50					55					60					
Met	Val	Gln	Asn	Arg	Asp	Glu	Val	Ile	Arg	Val	Leu	Glu	Asn	Ser	Ser
65					70					75				80	
Ile	Gly	Leu	Cys	Leu	Asp	Thr	Gly	His	Leu	Ala	Cys	Gly	Gly	Thr	Asp
				85					90					95	
Val	Val	Glu	Leu	Val	Arg	Lys	Tyr	Ala	Asn	Arg	Val	Asp	Ile	Val	His
			100						105					110	
Ala	Lys	Asp	Val	His	Lys	Glu	Met	Ala	Asp	Lys	Leu	Leu	Pro	Gly	Glu
		115					120					125			
Ile	Thr	Trp	Ser	Glu	Gly	Ile	Arg	Ala	Gly	Met	Phe	Ala	Pro	Ile	Gly
	130					135					140				
Asp	Gly	Asp	Ile	Asp	Phe	Ala	Ala	Ile	Val	Arg	Leu	Leu	Asp	Glu	Ala
145					150					155				160	
Gly	Phe	Asp	Gly	Tyr	Tyr	Val	Leu	Glu	Gln	Asp	Ile	Met			
				165						170					

<210> 103

<211> 321

<212> DNA

<213> Homo sapiens

<400> 103

nnccatgggg gaagacaaca gccatgtggg ggagaccgga gccattgggg ggagaccctt
 60
 gccattgggg ggagaccctt gccgtgggga aagacccttg ccatggggga gacccttgcc
 120
 actgggggga gacccttgcc gctgggggga gaccgagcc attgggggga gacccttgcc
 180
 atggggaaag acccctgcca ttgggggaga ntacctgcca ttgggggaga tccctgccgt
 240
 tggggggaga tccctgctgt tggggggaga ntccctcttg taggggaaga cccctgcagg
 300

agtgggtggg gcgaagaccc c
321

<210> 104
<211> 107
<212> PRT
<213> Homo sapiens

<400> 104
Xaa His Gly Gly Arg Gln Gln Pro Cys Gly Gly Asp Pro Ser His Trp
1 5 10 15
Gly Glu Thr Pro Ala Ile Gly Gly Arg Pro Leu Pro Trp Gly Lys Thr
20 25 30
Pro Ala Met Gly Gln Thr Pro Ala Thr Gly Gly Arg Pro Leu Pro Leu
35 40 45
Gly Gly Asp Pro Ser His Trp Gly Glu Thr Pro Ala Met Gly Lys Asp
50 55 60
Pro Cys His Trp Gly Arg Xaa Pro Ala Ile Gly Gly Asp Pro Cys Arg
65 70 75 80
Trp Gly Glu Ile Pro Ala Val Gly Gly Arg Xaa Pro Pro Val Gly Glu
85 90 95
Asp Pro Cys Arg Ser Gly Trp Gly Glu Asp Pro
100 105

<210> 105
<211> 344
<212> DNA
<213> Homo sapiens

<400> 105
nnntctctcc gaccgcgtcc agatccaccg tggcccgcat gaaccagtcg ttgttgccctt
60
ccgggtcaac gaggggttgg cgcacggtcc actccgtggc gcccggggtg atgtgcaaca
120
gggcggggccc gcgcgcggcc gggcctgatt ccagcctctc gtgctcgtcc cagtacccat
180
ccagcgcatt cccccagcgg tcggcatccc agccgtggtc gccgtcgagc gccccaggg
240
cctcaatgtc gtcacggcg gccagttcca cccggcgga catctcgttg cggaccatga
300
cccgaaggc gcgggaattc tcggtcagtt tcggcggtgc cggc
344

<210> 106
<211> 62
<212> PRT
<213> Homo sapiens

<400> 106
Cys Ala Thr Gly Arg Ala Arg Ala Arg Pro Gly Leu Ile Pro Ala Ser
1 5 10 15
Arg Ala Arg Pro Ser Thr His Pro Ala His Arg Pro Ser Gly Arg His
20 25 30
Pro Ser Arg Gly Arg Arg Ala Pro Pro Gly Pro Gln Cys Arg His

35 40 45
 Arg Arg Pro Val Pro Pro Gly Gly Thr Ser Arg Cys Gly Pro
 50 55 60

<210> 107
 <211> 549
 <212> DNA
 <213> Homo sapiens

<400> 107
 acgcgttcct cggacgctgc tagtttgagc gcacaggcat tgggcggggc tcacggggga
 60
 gccgcttaat aaccgaccaa catgaaactc aagggtgcc ccttcctagc ggggaccctg
 120
 cacagaccgc aaaataagggt gttttgctct gccctcctca gttcacgtgg gcaccttgga
 180
 aactgaaga aggcattttc cgaactcact gtcctacgga cttattctcc gcactgtttt
 240
 cgctccttc gccctgttct cgtgactgac aggagcaggg gtcacaagca ggcagcccga
 300
 gagctctgct cacctggaaa agcatttttg tgtagcttaa atgtgaaggc ctcaggcagt
 360
 ggctgttgt cctcctccac atgcgcccat cttactctt tcatgtgact ggctgtttt
 420
 tgaaggcaag gcccctgtca cccttggtta ggccaggtat gttctgcacc gaaaatggcc
 480
 ctgccctctg cattggatgg ctagctctta ggttggttta ttttagcaaa taagcggtac
 540
 agggtaggc
 549

<210> 108
 <211> 108
 <212> PRT
 <213> Homo sapiens

<400> 108
 Met Lys Leu Lys Gly Cys Pro Phe Leu Ala Gly Thr Leu His Arg Pro
 1 5 10 15
 Glu Asn Lys Gly Phe Cys Ser Ala Leu Ser Ser Arg Gly His Leu
 20 25 30
 Gly Thr Leu Lys Lys Ala Phe Ser Glu Leu Thr Val Leu Arg Thr Tyr
 35 40 45
 Ser Pro His Cys Phe Arg Leu Leu Arg Pro Val Leu Val Thr Asp Arg
 50 55 60
 Ser Arg Gly His Lys Gln Ala Ala Arg Glu Leu Cys Ser Pro Gly Lys
 65 70 75 80
 Ala Phe Leu Cys Ser Leu Asn Val Lys Ala Ser Gly Ser Gly Leu Leu
 85 90 95
 Ser Ser Ser Thr Cys Ala His Leu His Ser Phe Met
 100 105

<210> 109
 <211> 748

<212> DNA

<213> Homo sapiens

<400> 109

nngaattcag atttactttt tgcatttcct tgaatataaa ataggcagta aactaaccta
 60
 aagtcagaag acttatttgt tctggtacct attctgctgt taataaattg ataaatgagt
 120
 gatattgggg taagttacca atttacttta cagcccttaa gtaaataatc tgctttcctc
 180
 agcatcatag acttttgaag aggattaatt aagcgcttaa aaaacctgta gactctatta
 240
 cagtcagtga aaggaataat tctctttaca aagtaaatgc agttgtttta ttttagacaa
 300
 gagtgttcta aacttcgtga agagttaagg cttcaacatg aagaggataa gaagtcagca
 360
 atgtctcaac ttttgagtt gaaagatcga gagaaaaatg cagcaagaga ttcattggcag
 420
 aagaagtag aagatctctt aaaccagatt tccttgctga aacagaatct ggagatacag
 480
 ctttcccagt ctcagacttc tttgcaacaa ctgcaagccc agtttacgca agaacgacag
 540
 cggcttacgc aagagcttga agaattagag gagcaacatc agcaaagaca caaatcatta
 600
 aaagaagcac atgtccttgc atttcaaact atggaagagg aaaaggaaaa ggagcaaaga
 660
 gctcttgaaa atcatttaca acagaagcat tctgcagagc ttcaatcact aaaagatgca
 720
 cacagagagt caatggaggg cttccgga
 748

<210> 110

<211> 157

<212> PRT

<213> Homo sapiens

<400> 110

Met	Gln	Leu	Phe	Tyr	Phe	Arg	Gln	Glu	Cys	Ser	Lys	Leu	Arg	Glu	Glu
1			5						10					15	
Leu	Arg	Leu	Gln	His	Glu	Glu	Asp	Lys	Lys	Ser	Ala	Met	Ser	Gln	Leu
			20					25					30		
Leu	Gln	Leu	Lys	Asp	Arg	Glu	Lys	Asn	Ala	Ala	Arg	Asp	Ser	Trp	Gln
			35				40					45			
Lys	Lys	Val	Glu	Asp	Leu	Leu	Asn	Gln	Ile	Ser	Leu	Leu	Lys	Gln	Asn
			50				55				60				
Leu	Glu	Ile	Gln	Leu	Ser	Gln	Ser	Gln	Thr	Ser	Leu	Gln	Gln	Leu	Gln
						70				75				80	
Ala	Gln	Phe	Thr	Gln	Glu	Arg	Gln	Arg	Leu	Thr	Gln	Glu	Leu	Glu	Glu
						85			90					95	
Leu	Glu	Glu	Gln	His	Gln	Gln	Arg	His	Lys	Ser	Leu	Lys	Glu	Ala	His
						100			105				110		
Val	Leu	Ala	Phe	Gln	Thr	Met	Glu	Glu	Glu	Lys	Glu	Lys	Glu	Gln	Arg
						115			120				125		
Ala	Leu	Glu	Asn	His	Leu	Gln	Gln	Lys	His	Ser	Ala	Glu	Leu	Gln	Ser

130 135 140
 Leu Lys Asp Ala His Arg Glu Ser Met Glu Gly Phe Arg
 145 150 155

<210> 111
 <211> 429
 <212> DNA
 <213> Homo sapiens

<400> 111
 gcgcgcccag agagtgcaca atggtgtcag gatatgggag ccaccggaat cattaaccat
 60
 cgtgaaccgc tcgccccgca ggtcaacgac ttcgggatca ccgggttcga cggcattctc
 120
 tcgggttatn nacgccacca gcatncgact ttggctgaga tcatcgacc gttcggacat
 180
 ctggtcatga tcgacggaac cgactcattc gatctcatgg ccttcaagtc aaagtcgtta
 240
 acggtgacca gcgagtcgat gttcagccgt ccacagttcg cgacgcccga cgtcgccgaa
 300
 caaggccggg cactggccag catcgccgac ctcgctgaga aggggcagat cgtccgacg
 360
 atgaccgccc atatcgaggg tctgacaacc cagcatgtgc gtgaggccac cgcagccgtc
 420
 gagtccggc
 429

<210> 112
 <211> 143
 <212> PRT
 <213> Homo sapiens

<400> 112
 Ala Arg Pro Glu Ser Ala Gln Trp Cys Gln Asp Met Gly Ala Thr Gly
 1 5 10 15
 Ile Ile Asn His Arg Glu Pro Leu Ala Pro Gln Val Asn Asp Phe Gly
 20 25 30
 Ile Thr Gly Phe Asp Gly Ile Leu Ser Ala Tyr Xaa Arg His Gln His
 35 40 45
 Xaa Thr Leu Ala Glu Ile Ile Ala Pro Phe Gly His Leu Val Met Ile
 50 55 60
 Asp Gly Thr Asp Ser Phe Asp Leu Met Ala Phe Lys Ser Lys Ser Leu
 65 70 75 80
 Thr Val Thr Ser Glu Ser Met Phe Ser Arg Pro Gln Phe Ala Thr Pro
 85 90 95
 Asp Val Ala Glu Gln Gly Arg Ala Leu Ala Ser Ile Ala Asp Leu Val
 100 105 110
 Glu Lys Gly Gln Ile Arg Pro Thr Met Thr Arg His Ile Glu Gly Leu
 115 120 125
 Thr Thr Gln His Val Arg Glu Ala Thr Ala Ala Val Glu Ser Gly
 130 135 140

<210> 113
 <211> 382

<212> DNA

<213> Homo sapiens

<400> 113

gtgactgga cggatgctgg ggaactgagg cctgcccttc tccactcctg ctgcctctgt
60
gtggaccgcc tgccacctgg cttcaacgat gtggacgctc tgtgccgggc gctgtcagct
120
gtccacagcc ccaccttctg ccagctggcg tgcggccagg atgggcagct gaagggttc
180
gcggtgctgg agtatgagac ggctgagatg gcggaggagg cacagcagca ggcggacggc
240
ctgtccctgg ggggcagcca cctgcgagtc tccttctgcg cccctgggac ccccgggcgc
300
agtatgctgg ccgtctcat cgttgccag gccacggccc tcaatcgggg gcagggagtc
360
ctccccgagc ccaacatcct gc
382

<210> 114

<211> 125

<212> PRT

<213> Homo sapiens

<400> 114

Met	Leu	Gly	Ser	Gly	Arg	Thr	Pro	Cys	Pro	Arg	Leu	Arg	Ala	Val	Ala
1				5					10					15	
Trp	Ala	Thr	Met	Arg	Ala	Ala	Ser	Ile	Leu	Arg	Pro	Gly	Val	Pro	Gly
			20					25					30		
Ala	Gln	Lys	Glu	Thr	Arg	Arg	Trp	Leu	Pro	Pro	Arg	Asp	Arg	Pro	Ser
		35					40					45			
Ala	Cys	Cys	Cys	Ala	Ser	Ser	Ala	Ile	Ser	Ala	Val	Ser	Tyr	Ser	Ser
	50					55					60				
Thr	Ala	Lys	Pro	Phe	Ser	Cys	Pro	Ser	Trp	Pro	His	Ala	Ser	Trp	Gln
65					70				75					80	
Lys	Val	Gly	Leu	Trp	Thr	Ala	Asp	Ser	Ala	Arg	His	Arg	Ala	Ser	Thr
			85					90						95	
Ser	Leu	Lys	Pro	Gly	Gly	Arg	Arg	Ser	Thr	Gln	Arg	Gln	Gln	Glu	Trp
		100						105						110	
Arg	Arg	Ala	Gly	Leu	Ser	Ser	Pro	Ala	Ser	Val	Gln	Cys			
		115					120					125			

<210> 115

<211> 4798

<212> DNA

<213> Homo sapiens

<400> 115

ctcccttcca gtctgtctta agtactgatg ccagagaggt ctgcctaagc tcagtgtgtg
60
catttcactc cctgtttgga gccatgagtg gccccctgtt gccttcaaga ctgaagccat
120
ccccatcctt cctcccacca acctagaggc ttgtcttcgt aaatgctggc cctttccttc
180

atgggctcca ccctctgagt gtgtcatgaa ctcttccact tccttgcctt ggttcgtatt
240
gggtccctctg cccgagggtca gagatttgga cgagcccttc tcctccatct tcacagtctc
300
catttaccat tagatgccag gccagacag ttgatgaaaa atgtaagaaa cggacccaag
360
tataagaaaa agctagatct gtccatttta tttctagttt ttggaagaga tgtggataca
420
gagaaatatt actctactat tgaaaaaat atatatcagt ctgattataa atgtcaactc
480
atactaagtc aatatcaggg aagcagtgtg gagaaggatg tggcaaactg gagattgctg
540
tttccataga aagagggcag ctaatgctta tctccagacc attgttatgc agtaatgtag
600
actccatttc accagattta attcttaagg agaaactgga aacttgtatt attatttgat
660
tattactttt ggcagttaat tttaaaaatc accgtttagg caaaacaac atgacatttg
720
tctgggctac aagggactct tagcgactct tagcttgac ttacctttct aatctcaggg
780
ctttttacaa agcccttctc tgaattccct gaagatatgc tagcttgaag gctcactgca
840
tactctcaac ttcccagcca ctccctgtt tcctcttat accagttaat actttgatgt
900
tattttttca gtcctattta atcttttaga ctggagggtc cttgaggggtg gggcctaatt
960
cttggtttta tcttcccat agcaattatc aagttatctg ctccctttgg cacctaagaa
1020
atattagttg gggtgataac tgaatggctg tccttctgcc ttcgttactg cttttgtgga
1080
tgttattttc ttgacccct attccaccog tgtgcagtgg agaggatggg tgttgatatt
1140
tacagaggaa ggaattaact gagacgcaac aggatgatgg gaacttctca aggtcatgtg
1200
gcaaggaagt caaggaattg gggtttgaac ccaagtcgtc tcagttccat tccattgtcc
1260
tctacccct gccacctctc accttcttcc ctttcacctt tttctgtagc tgaacggaaa
1320
ccaccgtct ttaacatgaa tgcgatgagt gccttatacc acattgcaca gaacgaatcc
1380
cccgtgctcc agtcaggaca ctggctgag tacttccgga attttgtcga ctctgtctt
1440
cagaaaatcc ctcaagacag accaacctca gaggttctcc tgaagcaccg ctttggtgctc
1500
cgggagcggc caccacagt catcatggac ctgatccaga ggaccaagga tgccgtgcgg
1560
gagctggaca acctgcagta ccgca_gatg aagaagatcc tgttccaaga ggcaccaac
1620
ggccctgggt ccgaggcccc agaggaggaa gaggaggccg agccctacat gcaccgggccc
1680
gggactctga ccagcctcga gagtagccac tcagtgccca gcatgtccat cagcgcctcc
1740
agccagagca gctccgtcaa cagcctagca gatgcctcag acaacgagga agaggaggag
1800

gaggaggagg aagaggagga ggaggaagaa ggcctgaag cccgggagat ggccatgatg
1860
caggaggggg agcacacagt cacctctcac agctccatta tccaccggct gccgggctct
1920
gacaacctat atgatgaccc ctaccagcca gagataaccc ccagccctct ccagccgcct
1980
gcagccccag ctcccacttc caccacctct tccgcccgcc gccgggccta ctgccgtaac
2040
cgagaccact ttgccaccat ccgaaccgcc tccctggta gccgtcagat ccaggagcat
2100
gagcaggact ctgcgctgcg ggagcagctg agcggctata agcggatgcg acgacagcac
2160
cagaagcagc tgctggccct ggagtacagg ctgaggggtg aacgggagga gcacagtga
2220
cggctgcagc gggagcttga ggcgcagcgg gctggctttg gggcagaggc agaaaagctg
2280
gcccggcggc accaggccat aggtgagaag gaggcacgag ctgcccaggc cgaggagcgg
2340
aagttccagc agcacatcct tgggcagcag aagaaggagc tggctgccct gctggaggca
2400
cagaagcggg cctacaaact tcgcaaggaa cagctgaagg aggagctcca ggagaacccc
2460
agcactccca agcgggagaa ggcgagtggt ctgctgcggc agaaggagca gctccagcag
2520
tgccaggcgg agggaggaag agggctgctg cggcggcagc gccagtactt tgagctgcag
2580
tgtcgccagt acaagcgcaa gatgttgctg gctcggcaca gcctggacca ggacctgctg
2640
cgggaggacc tgaacaagaa gcagaccagc aaggacttgg agtgtgcaact gctgcttcgg
2700
cagcagagg ccacgcggga gctggagctg cggcagctcc aggcctgca gcgcacgcgg
2760
gctgagctca cccgcctgca gcaccagacg gagctgggca accagctgga gtacaacaag
2820
cggcgtgagc aagagttgcg gcagaagcat gcggcccagg ttccgagca gcccaagagc
2880
ctcaaagtac gtgcaggcca gcgcccccg ggccttcac tccccattcc tggggctctg
2940
ggcccaccca acacaggcac ccctatagaa cagcagccct gctcacctgg ccaggaggca
3000
gtcctggacc aaagaatgct tggcgaggag gaggaagcag ttggagagag aaggattctg
3060
ggaaaggaag gggccacttt ggagcccaag cagcagagga ttctggggga agaatacagga
3120
gcccctagtc ccagtccaca aaaacatggg agcctgggtg atgaggaagt ttggggctctg
3180
cctgaggaga tagaggagct tagggtgccc tccctgtac ccaggagag gagcattggt
3240
ggccaggagg aggctgggac gtggagcttg tgggggaagg aggatgagag tcttctggat
3300
gaggagtttg agcttgctg ggtccagggc ccagcactga ctcccgtccc tgaggaggag
3360
gaagaagagg aagagggggc tccgattggg acccctaggg atcctggaga tggttgtcct
3420

tcccccgaca tccctcctga accccctcca acacacctga ggccctgccc tgccagccag
3480
ctccctggac tectgtccca tggcctcctg gccggcctct cctttgcagt ggggtcctcc
3540
tctggcctcc tgccctcct gctgctgctg ctgcttccat tgctggcagc ccaggggtggg
3600
gggtggcctgc aggcagcgct gctggccctt gaggtggggc tgggtgggtct gggggcctcc
3660
tacctgctcc tttgtacagc cctgcacctg ccctccagtc ttttctact cctggcccag
3720
ggtaccgcac tgggggccgt cctgggcctg agctggcgcc gaggcctcat ggggtgtccc
3780
ctgggccttg gagctgcctg gctcttagct tggccaggcc tagctctacc tctggtggct
3840
atggcagcgg ggggcagatg ggtgcggcag caaggccccc ggggtgcggc gggcatactt
3900
cgactctggt tgcgggttct gctgcgcctg tcacccatgg ccttcggggc cctgcagggc
3960
tgtggggctg tgggggaccg ggggctgttt gcactgtacc ccaaaaccaa caaggatggc
4020
ttccgcagcc gcctgccgt ccctgggccc cggcggcgta atccccgcac cacccaacac
4080
ccattagctc tgttggaag ggtctgggtc ctgtgcaagg gctggaactg gcgtctggca
4140
cgggccagcc agggtttagc atcccacttg ccccgctggg ccatccacac actggccagc
4200
tggggcctgc ttcggggtga acggcccacc cgaatcccc ggctactacc acgcagccag
4260
cgccagctgg ggccccctgc ctcccgccag cactgccag ggactctagc cgggcggagg
4320
tcacgcaccc gccagtcccg ggccctgccc ccctggaggt agctgactcc agcccttcca
4380
gccccaatct agagcattga gcactttatc tcccacgact cagtgaagtt tctccagtcc
4440
ctagtctct cttttcaccc accttctca gtttgcac ttaccccagg cccagccctt
4500
cggacctcta gacaggcagc ctctcagct gtggagtcca gcagtcactc tgtgttctcc
4560
tggcgctct cccctaagtt attgctgttc gccgctgtg tgtgctcatc ctcacctca
4620
ttgactcagg cctggggcca ggggtgggtg aggggtgggaa gagtcatgtt ttttttctcc
4680
tctttgatct tgttttctg tctcccttc aacctgtccc cttcccccca ccaaaaaaag
4740
aaaaagacaa acacaaataa aatatctgag cggaactgtg aaaaaaaaaa aaaaaaag
4798

<210> 116

<211> 1062

<212> PRT

<213> Homo sapiens

<400> 116

Met Met Gly Thr Ser Gln Gly His Val Ala Arg Lys Ser Arg Asn Trp

1 5 10 15
 Gly Leu Asn Pro Ser Arg Leu Ser Ser Ile Pro Leu Ser Ser Thr Pro
 20 25 30
 Cys His Leu Ser Pro Ser Ser Leu Ser Pro Phe Ser Val Ala Glu Arg
 35 40 45
 Lys Pro Pro Leu Phe Asn Met Asn Ala Met Ser Ala Leu Tyr His Ile
 50 55 60
 Ala Gln Asn Glu Ser Pro Val Leu Gln Ser Gly His Trp Ser Glu Tyr
 65 70 75 80
 Phe Arg Asn Phe Val Asp Ser Cys Leu Gln Lys Ile Pro Gln Asp Arg
 85 90 95
 Pro Thr Ser Glu Val Leu Leu Lys His Arg Phe Val Leu Arg Glu Arg
 100 105 110
 Pro Pro Thr Val Ile Met Asp Leu Ile Gln Arg Thr Lys Asp Ala Val
 115 120 125
 Arg Glu Leu Asp Asn Leu Gln Tyr Arg Lys Met Lys Lys Ile Leu Phe
 130 135 140
 Gln Glu Ala Pro Asn Gly Pro Gly Ala Glu Ala Pro Glu Glu Glu Glu
 145 150 155 160
 Glu Ala Glu Pro Tyr Met His Arg Ala Gly Thr Leu Thr Ser Leu Glu
 165 170 175
 Ser Ser His Ser Val Pro Ser Met Ser Ile Ser Ala Ser Ser Gln Ser
 180 185 190
 Ser Ser Val Asn Ser Leu Ala Asp Ala Ser Asp Asn Glu Glu Glu Glu
 195 200 205
 Glu Glu Glu Glu Glu Glu Glu Glu Glu Gly Pro Glu Ala Arg
 210 215 220
 Glu Met Ala Met Met Gln Glu Gly Glu His Thr Val Thr Ser His Ser
 225 230 235 240
 Ser Ile Ile His Arg Leu Pro Gly Ser Asp Asn Leu Tyr Asp Asp Pro
 245 250 255
 Tyr Gln Pro Glu Ile Thr Pro Ser Pro Leu Gln Pro Pro Ala Ala Pro
 260 265 270
 Ala Pro Thr Ser Thr Thr Ser Ser Ala Arg Arg Arg Ala Tyr Cys Arg
 275 280 285
 Asn Arg Asp His Phe Ala Thr Ile Arg Thr Ala Ser Leu Val Ser Arg
 290 295 300
 Gln Ile Gln Glu His Glu Gln Asp Ser Ala Leu Arg Glu Gln Leu Ser
 305 310 315 320
 Gly Tyr Lys Arg Met Arg Arg Gln His Gln Lys Gln Leu Leu Ala Leu
 325 330 335
 Glu Ser Arg Leu Arg Gly Glu Arg Glu Glu His Ser Ala Arg Leu Gln
 340 345 350
 Arg Glu Leu Glu Ala Gln Arg Ala Gly Phe Gly Ala Glu Ala Glu Lys
 355 360 365
 Leu Ala Arg Arg His Gln Ala Ile Gly Glu Lys Glu Ala Arg Ala Ala
 370 375 380
 Gln Ala Glu Glu Arg Lys Phe Gln Gln His Ile Leu Gly Gln Gln Lys
 385 390 395 400
 Lys Glu Leu Ala Ala Leu Leu Glu Ala Gln Lys Arg Thr Tyr Lys Leu
 405 410 415
 Arg Lys Glu Gln Leu Lys Glu Glu Leu Gln Glu Asn Pro Ser Thr Pro
 420 425 430
 Lys Arg Glu Lys Ala Glu Trp Leu Leu Arg Gln Lys Glu Gln Leu Gln

435 440 445
 Gln Cys Gln Ala Glu Glu Glu Ala Gly Leu Leu Arg Arg Gln Arg Gln
 450 455 460
 Tyr Phe Glu Leu Gln Cys Arg Gln Tyr Lys Arg Lys Met Leu Leu Ala
 465 470 475 480
 Arg His Ser Leu Asp Gln Asp Leu Leu Arg Glu Asp Leu Asn Lys Lys
 485 490 495
 Gln Thr Gln Lys Asp Leu Glu Cys Ala Leu Leu Leu Arg Gln His Glu
 500 505 510
 Ala Thr Arg Glu Leu Glu Leu Arg Gln Leu Gln Ala Val Gln Arg Thr
 515 520 525
 Arg Ala Glu Leu Thr Arg Leu Gln His Gln Thr Glu Leu Gly Asn Gln
 530 535 540
 Leu Glu Tyr Asn Lys Arg Arg Glu Gln Glu Leu Arg Gln Lys His Ala
 545 550 555 560
 Ala Gln Val Arg Gln Gln Pro Lys Ser Leu Lys Val Arg Ala Gly Gln
 565 570 575
 Arg Pro Pro Gly Leu Pro Leu Pro Ile Pro Gly Ala Leu Gly Pro Pro
 580 585 590
 Asn Thr Gly Thr Pro Ile Glu Gln Gln Pro Cys Ser Pro Gly Gln Glu
 595 600 605
 Ala Val Leu Asp Gln Arg Met Leu Gly Glu Glu Glu Glu Ala Val Gly
 610 615 620
 Glu Arg Arg Ile Leu Gly Lys Glu Gly Ala Thr Leu Glu Pro Lys Gln
 625 630 635 640
 Gln Arg Ile Leu Gly Glu Glu Ser Gly Ala Pro Ser Pro Ser Pro Gln
 645 650 655
 Lys His Gly Ser Leu Val Asp Glu Glu Val Trp Gly Leu Pro Glu Glu
 660 665 670
 Ile Glu Glu Leu Arg Val Pro Ser Leu Val Pro Gln Glu Arg Ser Ile
 675 680 685
 Val Gly Gln Glu Glu Ala Gly Thr Trp Ser Leu Trp Gly Lys Glu Asp
 690 695 700
 Glu Ser Leu Leu Asp Glu Glu Phe Glu Leu Gly Trp Val Gln Gly Pro
 705 710 715 720
 Ala Leu Thr Pro Val Pro Glu Glu Glu Glu Glu Glu Glu Gly Ala
 725 730 735
 Pro Ile Gly Thr Pro Arg Asp Pro Gly Asp Gly Cys Pro Ser Pro Asp
 740 745 750
 Ile Pro Pro Glu Pro Pro Pro Thr His Leu Arg Pro Cys Pro Ala Ser
 755 760 765
 Gln Leu Pro Gly Leu Leu Ser His Gly Leu Leu Ala Gly Leu Ser Phe
 770 775 780
 Ala Val Gly Ser Ser Ser Gly Leu Leu Pro Leu Leu Leu Leu Leu
 785 790 795 800
 Leu Pro Leu Leu Ala Ala Gln Gly Gly Gly Leu Gln Ala Ala Leu
 805 810 815
 Leu Ala Leu Glu Val Gly Leu Val Gly Leu Gly Ala Ser Tyr Leu Leu
 820 825 830
 Leu Cys Thr Ala Leu His Leu Pro Ser Ser Leu Phe Leu Leu Leu Ala
 835 840 845
 Gln Gly Thr Ala Leu Gly Ala Val Leu Gly Leu Ser Trp Arg Arg Gly
 850 855 860
 Leu Met Gly Val Pro Leu Gly Leu Gly Ala Ala Trp Leu Leu Ala Trp

865 870 875 880
 Pro Gly Leu Ala Leu Pro Leu Val Ala Met Ala Ala Gly Gly Arg Trp
 885 890 895
 Val Arg Gln Gln Gly Pro Arg Val Arg Arg Gly Ile Ser Arg Leu Trp
 900 905 910
 Leu Arg Val Leu Leu Arg Leu Ser Pro Met Ala Phe Arg Ala Leu Gln
 915 920 925
 Gly Cys Gly Ala Val Gly Asp Arg Gly Leu Phe Ala Leu Tyr Pro Lys
 930 935 940
 Thr Asn Lys Asp Gly Phe Arg Ser Arg Leu Pro Val Pro Gly Pro Arg
 945 950 955 960
 Arg Arg Asn Pro Arg Thr Thr Gln His Pro Leu Ala Leu Leu Ala Arg
 965 970 975
 Val Trp Val Leu Cys Lys Gly Trp Asn Trp Arg Leu Ala Arg Ala Ser
 980 985 990
 Gln Gly Leu Ala Ser His Leu Pro Pro Trp Ala Ile His Thr Leu Ala
 995 1000 1005
 Ser Trp Gly Leu Leu Arg Gly Glu Arg Pro Thr Arg Ile Pro Arg Leu
 1010 1015 1020
 Leu Pro Arg Ser Gln Arg Gln Leu Gly Pro Pro Ala Ser Arg Gln Pro
 1025 1030 1035 1040
 Leu Pro Gly Thr Leu Ala Gly Arg Arg Ser Arg Thr Arg Gln Ser Arg
 1045 1050 1055
 Ala Leu Pro Pro Trp Arg
 1060

<210> 117

<211> 471

<212> DNA

<213> Homo sapiens

<400> 117

nacgcgttga cgatctgtct ggctgggtgta gtgatctgcg ctgtgggtgt cgtcgatgac
 60
 ctgctcgacc ttctgcctt ggccaaggca gctggccagg tattagcggc cggcatcgtc
 120
 gtcacgggcg gagtgcgaat gttttggatc ccgctgccga actccatcat tgctttgggg
 180
 acgcctactt cgatcttggg gacgggtgtt ttcattgtgt tgtgcgccaa tgcggtgaat
 240
 ttcattgatg gacttgacgg cctggcatcc ggtgtgggtg ccatcgggtc cttggctttc
 300
 ttctcataca cctacctgct ggctcacgaa caggactttg ttgttgcgac gactaccagt
 360
 ctcatcagg ctgcgacggc gggcgctgt ctcgggtttt tgccccacaa ctggcatccg
 420
 gcgaggatgt tcatgggtga ttccggagct ctgctacttg gcttattgct a
 471

<210> 118

<211> 157

<212> PRT

<213> Homo sapiens

<400> 118
 Xaa Ala Leu Thr Ile Cys Leu Ala Gly Val Val Ile Cys Ala Val Gly
 1 5 10 15
 Val Val Asp Asp Leu Leu Asp Leu Pro Ala Leu Ala Lys Ala Ala Gly
 20 25 30
 Gln Val Leu Ala Ala Gly Ile Val Val Thr Gly Gly Val Arg Met Phe
 35 40 45
 Trp Ile Pro Leu Pro Asn Ser Ile Ile Ala Leu Gly Thr Pro Thr Ser
 50 55 60
 Ile Leu Val Thr Val Phe Phe Ile Val Leu Cys Ala Asn Ala Val Asn
 65 70 75 80
 Phe Ile Asp Gly Leu Asp Gly Leu Ala Ser Gly Val Val Ala Ile Gly
 85 90 95
 Ser Leu Ala Phe Phe Ser Tyr Thr Tyr Leu Leu Ala His Glu Gln Asp
 100 105 110
 Phe Val Val Ala Thr Thr Thr Ser Leu Ile Thr Ala Ala Thr Ala Gly
 115 120 125
 Ala Cys Leu Gly Phe Leu Pro His Asn Trp His Pro Ala Arg Met Phe
 130 135 140
 Met Gly Asp Ser Gly Ala Leu Leu Leu Gly Leu Leu Leu
 145 150 155

<210> 119
 <211> 302
 <212> DNA
 <213> Homo sapiens

<400> 119
 ntcaaacatg agcagtcgtg gcggccgagg ccgcgggtggc tattatcgcg agctttatgg
 60
 tagccgaggt cgaggcagta aatctaataa aactttcgca aaaaattcgg atgtctactc
 120
 tcagaaaaag actcgaacag tacgaggcac ctccgaagat ttagcacgat cgctccataa
 180
 gcttcatatg cgcccgtagc ctgcgtatca tgacattgag ggtatgtggg ctttcccagc
 240
 ctttactttt tatctggatc atgcacaagc agaccatac gctgccccaa ataaggcacg
 300
 cn
 302

<210> 120
 <211> 98
 <212> PRT
 <213> Homo sapiens

<400> 120
 Met Ser Ser Arg Gly Gly Arg Gly Arg Gly Gly Tyr Tyr Arg Glu Leu
 1 5 10 15
 Tyr Gly Ser Arg Gly Arg Gly Ser Lys Ser Asn Glu Thr Phe Ala Lys
 20 25 30
 Asn Ser Asp Val Tyr Ser Gln Lys Lys Thr Arg Thr Val Arg Gly Thr
 35 40 45
 Ser Glu Asp Leu Ala Arg Ser Leu His Lys Leu His Met Arg Pro Tyr

50 55 60
 Pro Ala Tyr His Asp Ile Glu Gly Met Trp Ala Phe Pro Ala Phe Thr
 65 70 75 80
 Phe Tyr Leu Asp His Ala Gln Ala Asp Pro Tyr Ala Ala Pro Asn Lys
 85 90 95
 Ala Arg

<210> 121
 <211> 318
 <212> DNA
 <213> Homo sapiens

<400> 121
 ngcatggggg gccctgggac cgcacttggtg cccctttttt ttttagggaa aaaattgagc
 60
 cctaaaggat ttgccgcatt acaggaaagt tttttggtaa gtttgggggtt gtttctgtgc
 120
 tgtgtgagaa ggagtagaag cagctccagt agagtgggcc ttttcatttt tatccagagg
 180
 aaatttgtag gctgtggcta ttacttcctt ttttttcttt tttttttttg tttagagaca
 240
 gagtctgnct ctgtcgccag gctggagtga agtggcacga tctcagctca ctgcaacctc
 300
 tgcctcccag gttcaagc
 318

<210> 122
 <211> 89
 <212> PRT
 <213> Homo sapiens

<400> 122
 Xaa Met Gly Gly Pro Gly Thr Ala Leu Val Pro Leu Phe Phe Leu Gly
 1 5 10 15
 Lys Lys Leu Ser Pro Lys Gly Phe Ala Ala Leu Gln Glu Ser Phe Leu
 20 25 30
 Val Ser Leu Gly Leu Phe Leu Cys Cys Val Arg Arg Ser Arg Ser Ser
 35 40 45
 Ser Ser Arg Val Gly Leu Phe Ile Phe Ile Gln Arg Lys Phe Val Gly
 50 55 60
 Cys Gly Tyr Tyr Phe Leu Phe Phe Leu Phe Phe Cys Leu Glu Thr
 65 70 75 80
 Glu Ser Xaa Ser Val Ala Arg Leu Glu
 85

<210> 123
 <211> 338
 <212> DNA
 <213> Homo sapiens

<400> 123
 acgcgtctag ggtagaaatc aactccagta actgtcattc aacctcagca atgctgggga
 60

cgggcagagg cagggcagct gtgtgccaca ttctgccag ggctgggtcag gccccggctc
 120
 tcaccactcc tcttcctgc tttgaacctg tggaacaaag ggccccctgca ccccaactca
 180
 ttctcttttg ccacataagg gcctcaagtc atgctgtccc ctctgcctgg gttgcttttt
 240
 ctccctctgc ttgggtcact gttcacacca ctggccactt tcctcaggga agggccctca
 300
 ctgcccacac acctaaacat gccccctgct cctccata
 338

<210> 124
 <211> 96
 <212> PRT
 <213> Homo sapiens

<400> 124
 Met Leu Gly Thr Gly Arg Gly Arg Ala Ala Val Cys His Ile Pro Ala
 1 5 10 15
 Arg Ala Gly Gln Ala Pro Ala Leu Thr Thr Pro Pro Pro Cys Phe Glu
 20 25 30
 Pro Val Glu Gln Arg Ala Pro Ala Pro Gln Leu Ile Pro Leu Cys His
 35 40 45
 Ile Arg Ala Ser Ser His Ala Val Pro Ser Ala Trp Val Ala Phe Ser
 50 55 60
 Pro Ser Ala Trp Val Thr Val His Thr Thr Gly His Phe Pro Gln Gly
 65 70 75 80
 Arg Ala Leu Thr Ala His Thr Pro Lys His Ala Pro Cys Ser Ser Ile
 85 90 95

<210> 125
 <211> 280
 <212> DNA
 <213> Homo sapiens

<400> 125
 ccatggacct ggccagccac catcacctgc ctctgcctc acccaccctg ggtgcctgcc
 60
 ggcaaggatt ggagggcaga ctgctggagc gtgagaccag gccaatctgt ctttctggga
 120
 accttcagcc tccaactgga gctgactgtc aactttcggg tgagaagtca ctttctgca
 180
 ttcccaccac actatctatc tgtgcaatac ggcagcgtga cagcactcac cttattgagg
 240
 gcttctgctg tcttggecca ttctggatag gcctgatcta
 280

<210> 126
 <211> 92
 <212> PRT
 <213> Homo sapiens

<400> 126
 Met Asp Leu Ala Ser His His His Leu Pro Pro Ala Ser Pro Thr Leu

```

      1           5           10           15
Gly Ala Cys Arg Gln Gly Leu Glu Gly Arg Leu Leu Glu Arg Glu Thr
      20           25           30
Arg Pro Ile Cys Leu Ser Gly Asn Leu Gln Pro Pro Thr Gly Ala Asp
      35           40           45
Cys Gln Leu Ser Gly Glu Lys Ser Leu Phe Cys Ile Pro Thr Thr Leu
      50           55           60
Ser Ile Cys Ala Ile Arg Gln Arg Asp Ser Thr His Leu Ile Glu Gly
      65           70           75           80
Phe Cys Cys Pro Gly Pro Phe Trp Ile Gly Leu Ile
      85           90

```

<210> 127

<211> 444

<212> DNA

<213> Homo sapiens

<400> 127

```

cgcgatgatcg ccgtggcgga gggccgcggc gccgactcga tcgcccagct gacaaccgag
60
ctgcaaagcc gtcactgccc tgccggagcag atcacgtccg tcagcatcga catgtcgcca
120
gcgttcatca ggggctgctc cgagcacctg cccaacgcgc gcgtcacctt cgacaagttc
180
cacgtcatcg ggcacgcca tgcggccgtg gacaggatgc gccgcatcga gcagcgcagc
240
gacaagtccc tcaaggggat gcgctggctg ctgctgaaga accgcgccag cctcaagccc
300
gaggctgccg ccgatctgga tgccctgatc gccaggatgg ccactgtgctg caccgcgcgc
360
gcctgggtct acaaggagca gctgcgcgag atcctcgcgc gcaagcagat caacgtggca
420
cgcgacatgc tcaagcactg gtgc
444

```

<210> 128

<211> 148

<212> PRT

<213> Homo sapiens

<400> 128

```

Arg Val Ile Ala Val Ala Glu Gly Arg Gly Ala Asp Ser Ile Ala Gln
      1           5           10           15
Leu Thr Thr Glu Leu Gln Ser Arg His Cys Pro Ala Glu Gln Ile Thr
      20           25           30
Ser Val Ser Ile Asp Met Ser Pro Ala Phe Ile Arg Gly Cys Ala Glu
      35           40           45
His Leu Pro Asn Ala Arg Val Thr Phe Asp Lys Phe His Val Ile Gly
      50           55           60
His Ala Asn Ala Ala Val Asp Arg Met Arg Arg Ile Glu Gln Arg Ser
      65           70           75           80
Asp Lys Ser Leu Lys Gly Met Arg Trp Ser Leu Leu Lys Asn Arg Ala
      85           90           95
Ser Leu Lys Pro Glu Ala Ala Ala Asp Leu Asp Ala Leu Ile Ala Arg

```

100 105 110
 Met Ala Thr Val Arg Thr Ala Arg Ala Trp Val Tyr Lys Glu Gln Leu
 115 120 125
 Arg Glu Ile Leu Ala Arg Lys Gln Ile Asn Val Ala Arg Asp Met Leu
 130 135 140
 Lys His Trp Cys
 145

<210> 129
 <211> 291
 <212> DNA
 <213> Homo sapiens

<400> 129
 gaggagggac gtaccgtccc cggtatagcc aagctcgaga agccgcaagc tatcgagaac
 60
 ttggacgaga ttattgacgt ctttgacgcc gtcattggtg cccgtggcga tatggccgtc
 120
 gagtgtccgc tgcaggaagt tccgctgac caaaagcaga tcatcgagaa ggctcgttta
 180
 caggctaagc ccgtcattgt ggccaccag atgcttgagt cgatgatcca cgctccccgt
 240
 ccgacccgcg ctgagggcgc cgacgtcgcg aacgccatcc ttgacggcgc g
 291

<210> 130
 <211> 97
 <212> PRT
 <213> Homo sapiens

<400> 130
 Glu Glu Gly Arg Thr Val Pro Val Ile Ala Lys Leu Glu Lys Pro Gln
 1 5 10 15
 Ala Ile Glu Asn Leu Asp Glu Ile Ile Asp Val Phe Asp Ala Val Met
 20 25 30
 Val Ala Arg Gly Asp Met Ala Val Glu Cys Pro Leu Glu Val Pro
 35 40 45
 Leu Ile Gln Lys Gln Ile Ile Glu Lys Ala Arg Leu Gln Ala Lys Pro
 50 55 60
 Val Ile Val Ala Thr Gln Met Leu Glu Ser Met Ile His Ala Pro Arg
 65 70 75 80
 Pro Thr Arg Ala Glu Ala Ala Asp Val Ala Asn Ala Ile Leu Asp Gly
 85 90 95
 Ala

<210> 131
 <211> 416
 <212> DNA
 <213> Homo sapiens

<400> 131
 tccggagcgt ccgtggccct catgggtgtg tcagcgtggt tgctgtctcg ggccgcagag
 60

attccaccgg tgctctacct ggaggccgca gccgtcgggg ttcgattctt cggcattctcc
 120
 cgcggtgtct tccgctacgc cgaacgtctg gtaggccacg acctgggtct gcggatgcag
 180
 ggggcattgc ggatgcgggt ctacgaccgg ctgtcacgta ccnaccctgc tgggnnacgt
 240
 cgccgggggtg acctgctggt acgggttact gccgacgtcg acgcgggtgtt ggacatggtc
 300
 gtgcgggtga tcgttccggc gtgcgcgtca agcctcgtca tcattggcac caccgtcctt
 360
 ctttgtccga gagaagggtg agttttctta gccggattcc aacacagcct gggggc
 416

<210> 132

<211> 126

<212> PRT

<213> Homo sapiens

<400> 132

Ser	Gly	Ala	Ser	Val	Ala	Leu	Met	Gly	Val	Ser	Ala	Trp	Leu	Leu	Ser
1				5					10					15	
Arg	Ala	Ala	Glu	Ile	Pro	Pro	Val	Leu	Tyr	Leu	Glu	Ala	Ala	Ala	Val
			20					25					30		
Gly	Val	Arg	Phe	Phe	Gly	Ile	Ser	Arg	Gly	Val	Phe	Arg	Tyr	Ala	Glu
		35				40					45				
Arg	Leu	Val	Gly	His	Asp	Leu	Ala	Leu	Arg	Met	Gln	Gly	Ala	Leu	Arg
	50				55					60					
Met	Arg	Val	Tyr	Asp	Arg	Leu	Ser	Arg	Thr	Xaa	Pro	Ala	Gly	Xaa	Arg
65				70					75					80	
Arg	Arg	Gly	Asp	Leu	Leu	Val	Arg	Val	Thr	Ala	Asp	Val	Asp	Ala	Val
			85					90					95		
Leu	Asp	Met	Val	Val	Arg	Val	Ile	Val	Pro	Ala	Cys	Ala	Ser	Ser	Leu
		100					105					110			
Val	Ile	Ile	Gly	Thr	Thr	Val	Leu	Leu	Cys	Pro	Arg	Glu	Gly		
	115					120						125			

<210> 133

<211> 327

<212> DNA

<213> Homo sapiens

<400> 133

gccgttgcta tcgctgctgg tatgcgtgca gacgtcactg tttttgatat caatatcgct
 60
 gcgttgaaga gactcgccga catctaccag ggtcgtgttc acacagtagt atccacccgc
 120
 gccgaaattg cgaaggcgct agaaaccgct gacgttgtga tcggttctgt ccttattccg
 180
 ggtagtctta ccccgaagct tgttactacc gatatgggtg ctcacatgca gcctgggtct
 240
 gttcttattg atattgctat agaccaaggc ggctgcttcg aggattcgca cccaccact
 300
 tacgatgacc ccactttcac tgtgcac
 327

<210> 134
 <211> 109
 <212> PRT
 <213> Homo sapiens

<400> 134
 Ala Val Ala Ile Ala Ala Gly Met Arg Ala Asp Val Thr Val Phe Asp
 1 5 10 15
 Ile Asn Ile Ala Ala Leu Lys Arg Leu Ala Asp Ile Tyr Gln Gly Arg
 20 25 30
 Val His Thr Val Val Ser Thr Arg Ala Glu Ile Ala Lys Ala Leu Glu
 35 40 45
 Thr Ala Asp Val Val Ile Gly Ser Val Leu Ile Pro Gly Ser Ser Thr
 50 55 60
 Pro Lys Leu Val Thr Thr Asp Met Val Ala His Met Gln Pro Gly Ser
 65 70 75 80
 Val Leu Ile Asp Ile Ala Ile Asp Gln Gly Gly Cys Phe Glu Asp Ser
 85 90 95
 His Pro Thr Thr Tyr Asp Asp Pro Thr Phe Thr Val His
 100 105

<210> 135
 <211> 560
 <212> DNA
 <213> Homo sapiens

<400> 135
 taagatgtgg tcctgccctg ttctgaagg ggctgcagct ctgatggaaa atacagggat
 60
 ttacactcag ggctacagcc acggggggct gaggcccaag gctgcaatct cgggggaagg
 120
 ggaagtggc ttttctggt ggattggaaa catcctcttg gaggcaaaga ctttctctgg
 180
 atcttacaga cttcccgga ttttagatt agaatttgg gggcaaagga ggctgtcttg
 240
 ttttaaagca atgctacata gacacagtgg ggaagacctg gttcgacggc agataagcag
 300
 tgggtgatgg gcttgaggag gagagtcagg gcaaagtcta agactgagca gaaaggaatt
 360
 ccccatctc ccatggataa gtacgttcta gaacattctc tttgggtcta atactctgaa
 420
 atgacatctt gtcttcatgc tcgagagaga attacttcac tggctccact tggagtcca
 480
 gtgttcagac accaagcctg actgggaggg ttccgttttc ttaacacctt cccaccgccg
 540
 acttccaagt cccacgcgt
 560

<210> 136
 <211> 100
 <212> PRT
 <213> Homo sapiens

<400> 136

Met Trp Ser Cys Pro Val Pro Glu Gly Ala Ala Ala Leu Met Glu Asn
 1 5 10 15
 Thr Gly Ile Tyr Thr Gln Gly Tyr Ser His Gly Gly Leu Arg Pro Lys
 20 25 30
 Ala Ala Ile Ser Gly Glu Gly Glu Val Gly Phe Ser Trp Ile Gly
 35 40 45
 Asn Ile Leu Leu Glu Ala Lys Thr Phe Pro Gly Ser Tyr Arg Leu Pro
 50 55 60
 Gly Ile Phe Arg Leu Glu Tyr Trp Gly Gln Arg Arg Leu Ser Cys Phe
 65 70 75 80
 Lys Ala Met Leu His Arg His Ser Gly Glu Asp Leu Val Arg Arg Gln
 85 90 95
 Ile Ser Ser Gly
 100

<210> 137

<211> 429

<212> DNA

<213> Homo sapiens

<400> 137

accggttgga tggcctgcag gccaaagcgt tcttgcaaac tcagcaggcc ttcagcgcaa
 60
 gagggcaaaca gctggctcgcg cacctgcttg aggtccaccg attgcgcata gcccttgagc
 120
 aaggcgcgcc agttgggttt gtcggccact tggctgcgga acaggtcttc gacaaaaccg
 180
 gactgctggc gggctcgcaac gcgcatgata ggcagcgccct ggctggcgcc ctggctcgagc
 240
 cagcgcgctcg gcagttgggt ggcccgggtg ataccgacct tgatccccga cgaattggcc
 300
 aggtacacca catggctcggc catgcagaat gtttcgcccc agccgggatac acggcaagtg
 360
 ccggcgctcgt aatggcaacg ttcggggctc atgatgcaca ggtcacactg ggccagcttg
 420
 gtcattgccc
 429

<210> 138

<211> 141

<212> PRT

<213> Homo sapiens

<400> 138

Met Thr Lys Leu Ala Gln Cys Asp Leu Cys Ile Met Ser Pro Glu Arg
 1 5 10 15
 Cys His Tyr Asp Ala Gly Thr Cys Arg Asp Pro Gly Trp Gly Glu Thr
 20 25 30
 Phe Cys Met Thr Asp His Val Val Tyr Leu Ala Asn Ser Ser Gly Ile
 35 40 45
 Lys Val Gly Ile Thr Arg Ala Thr Gln Leu Pro Thr Arg Trp Leu Asp
 50 55 60
 Gln Gly Ala Ser Gln Ala Leu Pro Ile Met Arg Val Ala Thr Arg Gln

```

65          70          75          80
Gln Ser Gly Phe Val Glu Asp Leu Phe Arg Ser Gln Val Ala Asp Lys
          85          90          95
Thr Asn Trp Arg Ala Leu Leu Lys Gly Asp Ala Gln Ser Val Asp Leu
          100          105          110
Lys Gln Val Arg Asp Gln Leu Phe Ala Ser Cys Ala Glu Gly Leu Leu
          115          120          125
Ser Leu Gln Glu Arg Phe Gly Leu Gln Ala Ile Gln Pro
          130          135          140

```

<210> 139
 <211> 341
 <212> DNA
 <213> Homo sapiens

```

<400> 139
acgcgtcggtt tgaaggcttg atccgcacgt ccaattcgct ttgcgcdaat gcgccgcagc
60
ttgtgaacag cagaatcaag ccgctggtaa atcttctctgg gagcttcata ggcggggatg
120
ctacacgagc tggggagaca ctttgaaccc ggaattgtct gaataattct gtctcaaacc
180
tttgcagcct gtaacgactg agggttcgga tggaaaaaca catgctccag gatgggaccg
240
acggccactt caccgatctc ttcatagccc tggcgtttgt agaaatccag gtagcgcgaa
300
tcgccagcgt cgagcacgac gcctgatgag tgcgggtcat t
341

```

<210> 140
 <211> 113
 <212> PRT
 <213> Homo sapiens

```

<400> 140
Met Thr Arg Thr His Gln Ala Ser Cys Ser Thr Leu Ala Ile Arg Ala
1          5          10          15
Thr Trp Ile Ser Thr Asn Ala Arg Ala Met Lys Arg Ser Val Lys Trp
          20          25          30
Pro Ser Val Pro Ser Trp Ser Met Cys Phe Ser Ile Arg Thr Leu Ser
          35          40          45
Arg Tyr Arg Leu Gln Arg Phe Glu Thr Glu Leu Phe Arg Gln Phe Arg
          50          55          60
Val Gln Ser Val Ser Pro Ala Arg Val Ala Ser Pro Pro Met Lys Leu
65          70          75          80
Pro Gly Arg Phe Thr Ser Gly Leu Ile Leu Leu Phe Thr Ser Cys Gly
          85          90          95
Ala Leu Ala Gln Ser Glu Leu Asp Val Arg Ile Lys Pro Ser Asn Asp
          100          105          110
Ala

```

<210> 141
 <211> 324

<212> DNA

<213> Homo sapiens

<400> 141

gaattcctct tggatagctt cgggtaaatg ggtacagcaa atatcaggag cgcaaccgca
60
acctttactt actggtacat gaacaccatt tacattacag ctatcgctact caccaccacgt
120
catgtgaaca gacacataac tgaaagggtt ataaaccaca gtctcacggt acgtatgacc
180
gtcaactgtg aacaccgcta agtaatagcc tgcgggggct tgcatagaact cctttgacca
240
tgcgtaataa atacgtccgt cattagtcac acctgatggg gcgaaacaaa aagaacggca
300
gcagttatca ccgccatac gcgt
324

<210> 142

<211> 106

<212> PRT

<213> Homo sapiens

<400> 142

Met	Gly	Gly	Asp	Asn	Cys	Cys	Arg	Ser	Phe	Cys	Phe	Ala	Pro	Ser	Gly
1				5					10					15	
Val	Thr	Asn	Asp	Gly	Arg	Ile	Tyr	Tyr	Ala	Trp	Ser	Lys	Glu	Phe	Met
		20					25					30			
Gln	Ala	Pro	Ala	Gly	Tyr	Tyr	Leu	Ala	Val	Phe	Thr	Val	Asp	Gly	His
		35				40					45				
Thr	Tyr	Arg	Glu	Thr	Val	Val	Tyr	Lys	Pro	Phe	Ser	Tyr	Val	Ser	Val
	50				55					60					
His	Met	Thr	Trp	Gly	Glu	Tyr	Asp	Ser	Cys	Asn	Val	Asn	Gly	Val	His
65				70					75				80		
Val	Pro	Val	Ser	Lys	Gly	Cys	Gly	Cys	Ala	Pro	Asp	Ile	Cys	Cys	Thr
			85				90					95			
His	Leu	Pro	Glu	Ala	Ile	Gln	Glu	Glu	Phe						
			100				105								

<210> 143

<211> 1325

<212> DNA

<213> Homo sapiens

<400> 143

nacgcgtgga tctgccagct gagcctggag ctgtgcaggc agctgccctg ctacgatgag
60
gcaccccagg agaagaactt cctgtacaaa tgcataaggca ccaccctggg tgctgcttca
120
agtaaggagg tggtaggaa gcaccttcaa gagctgctgg agacggccag ataccaggag
180
gaggcagaac gcgagggcct cgcctgctgc ttcgggatct gtgccatctc ccacctcgag
240
gacacgctgg ccagctgga ggacttcgtg aggtcagagg tcttcagaaa atccattggc
300

attctcaaca tttttaagga tcgaagtgag aacgaagtgg agaaggtgaa gagtgtctctg
 360
 atcctgtgct atgggcacgt ggcggcccg gccccccggg agctggtgct ggccaaggta
 420
 gagtcagaca tcctccggaa catcntgcca gcacttcagc acnncaagga cccagccctg
 480
 aagctgtgcc ttgtccagag tgtgtgcatg gtcagccgag ccatctgcag cagcaccag
 540
 gctggtcctt tccacttcac ccggaaagca gagctggtgg cacagatgat ggagtccatc
 600
 agggcagagc ccccggaacc cttgaggaca cctattcgga agaaagccat gctcacctgc
 660
 acttacttgg tctccgtgga gccagcgctg gacgagcagg cccgggcgga tgtgatccat
 720
 ggctgcctgc acagcatcat ggccctgctg cctgagccca aggaggagga cggaggctgc
 780
 cagaagtccc tgtatctgga gacactgcac gcccttgagg atctgtgcac gagectcctg
 840
 cagcggaaca tgacccccca aggcctgcag atcatgattg agcacctgag cccatggatc
 900
 aagtccccaa gaggtcacgt agcggcgctg gccctaggcc tgagcgccct cctcgtgcgc
 960
 tacttcctgg agcacctgcg tgtcagtggc gcccaagtag ataccaggtt tccatctgag
 1020
 cccaggatcc tgtgcaatgg ccttgggtgcc cttccacaac ctgggccttc tcatcgccct
 1080
 cttctcccca cgggtgtgagg acctgtggcc tgccaccgag caggaggccg tggactgtgt
 1140
 ctactccctg ctgtacctcc agctcggcta tgagggtctc tcccgggact accgcgatga
 1200
 cgtggcggag eggctcctca gcctcaagga cggcctcgtg caccctgacc ccgccattct
 1260
 cttccacacc tgccacagtg taggccagat tattgccaag cgctccccc cagcccttca
 1320
 cgcgt
 1325

<210> 144

<211> 390

<212> PRT

<213> Homo sapiens

<400> 144

Xaa Ala Trp Ile Cys Gln Leu Ser Leu Glu Leu Cys Arg Gln Leu Pro
 1 5 10 15
 Cys Tyr Asp Glu Ala Pro Gln Glu Lys Asn Phe Leu Tyr Lys Cys Ile
 20 25 30
 Gly Thr Thr Leu Gly Ala Ala Ser Ser Lys Glu Val Val Arg Lys His
 35 40 45
 Leu Gln Glu Leu Leu Glu Thr Ala Arg Tyr Gln Glu Glu Ala Glu Arg
 50 55 60
 Glu Gly Leu Ala Cys Cys Phe Gly Ile Cys Ala Ile Ser His Leu Glu
 65 70 75 80
 Asp Thr Leu Ala Gln Leu Glu Asp Phe Val Arg Ser Glu Val Phe Arg

85										90					95				
Lys	Ser	Ile	Gly	Ile	Leu	Asn	Ile	Phe	Lys	Asp	Arg	Ser	Glu	Asn	Glu				
100				105				110											
Val	Glu	Lys	Val	Lys	Ser	Ala	Leu	Ile	Leu	Cys	Tyr	Gly	His	Val	Ala				
115				120				125											
Ala	Arg	Ala	Pro	Arg	Glu	Leu	Val	Leu	Ala	Lys	Val	Glu	Ser	Asp	Ile				
130				135				140											
Leu	Arg	Asn	Ile	Xaa	Pro	Ala	Leu	Gln	His	Xaa	Lys	Asp	Pro	Ala	Leu				
145	150				155				160										
Lys	Leu	Cys	Leu	Val	Gln	Ser	Val	Cys	Met	Val	Ser	Arg	Ala	Ile	Cys				
165				170				175											
Ser	Ser	Thr	Gln	Ala	Gly	Ser	Phe	His	Phe	Thr	Arg	Lys	Ala	Glu	Leu				
180				185				190											
Val	Ala	Gln	Met	Met	Glu	Phe	Ile	Arg	Ala	Glu	Pro	Pro	Asp	Ser	Leu				
195				200				205											
Arg	Thr	Pro	Ile	Arg	Lys	Lys	Ala	Met	Leu	Thr	Cys	Thr	Tyr	Leu	Val				
210				215				220											
Ser	Val	Glu	Pro	Ala	Leu	Asp	Glu	Gln	Ala	Arg	Ala	Asp	Val	Ile	His				
225	230				235				240										
Gly	Cys	Leu	His	Ser	Ile	Met	Ala	Leu	Leu	Pro	Glu	Pro	Lys	Glu	Glu				
245				250				255											
Asp	Gly	Gly	Cys	Gln	Lys	Ser	Leu	Tyr	Leu	Glu	Thr	Leu	His	Ala	Leu				
260				265				270											
Glu	Asp	Leu	Leu	Thr	Ser	Leu	Leu	Gln	Arg	Asn	Met	Thr	Pro	Gln	Gly				
275				280				285											
Leu	Gln	Ile	Met	Ile	Glu	His	Leu	Ser	Pro	Trp	Ile	Lys	Ser	Pro	Arg				
290				295				300											
Gly	His	Val	Ala	Ala	Arg	Ala	Leu	Gly	Leu	Ser	Ala	Leu	Leu	Val	Arg				
305	310				315				320										
Tyr	Phe	Leu	Glu	His	Leu	Arg	Val	Ser	Gly	Ala	Gln	Val	Asp	Thr	Arg				
325				330				335											
Phe	Pro	Ser	Glu	Pro	Arg	Ile	Leu	Cys	Asn	Gly	Pro	Gly	Ala	Leu	Pro				
340				345				350											
Gln	Pro	Gly	Pro	Ser	His	Arg	Pro	Leu	Leu	Pro	Thr	Val	Cys	Gly	Pro				
355				360				365											
Val	Ala	Cys	His	Pro	Pro	Gly	Gly	Arg	Gly	Leu	Cys	Leu	Leu	Pro	Ala				
370				375				380											
Val	Pro	Pro	Ala	Arg	Leu														
385	390																		

<210> 145

<211> 802

<212> DNA

<213> Homo sapiens

<400> 145

cgcccgctcta ggtccggctc agtgcgctgt tgctcgccgt agaacacgag gctgcgcaag
60
cataagcaga cgtagagagt ggtcacatcc atgtcgatgg tgtgcgcgta atgaaggctc
120
acatcacctt ggtgaaggcc tgcaccacta gcgtcggcac catttccccg cgtcggacaa
180
gacatcatgc cccatatctt gacagaatgt ctgacatgag tatgccacgc cgagcagcac
240

cagaggacga caccgatctg gcggaacccg cccgttcattg ggcagatac ctcattcctcg
 300
 tcatttgtgg cgttatcgct gctgtcctcg gactaggcat ttccgggtat cttgcgtggg
 360
 ggtcattgtg cgatcaagct gccgggggtct gtcagcgtgg tgaaccggtt atgtactggg
 420
 gttcgggtgg ctctctggcc attctcggac tcattatcgg ggtcttgacg cagatctggc
 480
 tggagaagcg ctggtggcac atgcttgcca tcgtcatccc ggctgttttc atcgtcgccg
 540
 gtatcttttt ctggctcgcc gtctaagaag gggcgtcaca gattccacaa acgacacagg
 600
 tattgatctc cgttttatcg gctcctagca gccgtgggtca acgtatcgct atcaagcgat
 660
 acaggactcg tcgttcgcat cgttgttggt ctgctgggaa acaatcccag cgatctactc
 720
 ggctaccgcc agacagttca ctcacaaccc ctcacgccgg cgcagacatc aaatcccatt
 780
 ctcgatagac ggcccacacc ac
 802

<210> 146
 <211> 151
 <212> PRT
 <213> Homo sapiens

<400> 146
 Met Lys Val Tyr Ile Thr Leu Val Lys Ala Cys Thr Thr Ser Val Gly
 1 5 10 15
 Thr Ile Ser Pro Arg Arg Thr Arg His His Ala Pro Tyr Leu Asp Arg
 20 25 30
 Met Ser Asp Met Ser Met Pro Arg Arg Ala Ala Pro Glu Asp Asp Thr
 35 40 45
 Asp Leu Ala Asp Ala Ala Arg Ser Trp Arg Arg Tyr Leu Ile Leu Val
 50 55 60
 Ile Cys Gly Val Ile Val Ala Val Leu Gly Leu Gly Ile Phe Gly Tyr
 65 70 75 80
 Leu Ala Trp Trp Ser Leu Cys Asp Gln Ala Ala Gly Val Cys Gln Arg
 85 90 95
 Gly Glu Pro Val Met Tyr Trp Cys Ser Val Val Ser Leu Ala Ile Leu
 100 105 110
 Gly Leu Ile Ile Gly Val Leu Thr Gln Ile Trp Leu Glu Lys Arg Trp
 115 120 125
 Trp His Met Leu Ala Ile Val Ile Pro Ala Val Phe Ile Val Ala Gly
 130 135 140
 Ile Phe Phe Trp Leu Ala Val
 145 150

<210> 147
 <211> 368
 <212> DNA
 <213> Homo sapiens

<400> 147

acgcgtgaaa acggtatgac tcttctggcc ttagtagatc tgtctaaaaa acccgatgag
 60
 ttacacagct gggcattagt agcccgcat gttcatgaca ttcttggtct acgaaaagtt
 120
 attggtcaga aagtaccttg tgttgcatg acgggggtcgg aaaaggtgct tcataaaaag
 180
 gattactggg atctagcaac acctatgcca attgctggg gtacaacgga ccgaacagtt
 240
 attgctgatg cagcagctac aatccccacc acggagtggg atatccttgc aagactacgt
 300
 ccacgcctag aagaggttcg caagcaacgt aatgatgtat tgcctcctcaa cgaggaggat
 360
 ccccccta
 368

<210> 148

<211> 117

<212> PRT

<213> Homo sapiens

<400> 148

Met	Thr	Leu	Leu	Ala	Leu	Val	Asp	Leu	Ser	Lys	Lys	Pro	Asp	Glu	Phe
1				5						10				15	
Thr	Gln	Trp	Ala	Leu	Val	Ala	Arg	Asp	Val	His	Asp	Ile	Pro	Gly	Leu
			20					25					30		
Arg	Lys	Val	Ile	Gly	Gln	Lys	Val	Pro	Cys	Val	Ala	Val	Thr	Gly	Ser
		35					40					45			
Glu	Lys	Val	Leu	His	Lys	Lys	Asp	Tyr	Trp	Asp	Leu	Ala	Thr	Pro	Met
	50					55				60					
Pro	Ile	Ala	Trp	Gly	Thr	Thr	Asp	Arg	Thr	Val	Ile	Ala	Asp	Ala	Arg
65					70					75				80	
Arg	Thr	Ile	Pro	Thr	Thr	Glu	Trp	Asp	Ile	Leu	Ala	Arg	Leu	Arg	Pro
			85					90					95		
Arg	Leu	Glu	Glu	Val	Arg	Lys	Gln	Arg	Asn	Asp	Val	Leu	Leu	Leu	Asn
		100						105					110		
Glu	Glu	Asp	Pro	Pro											
		115													

<210> 149

<211> 407

<212> DNA

<213> Homo sapiens

<400> 149

nngctagcat ggaccctagt cacacaggca gccatacccg aggtcaaagt gacccatttt
 60
 cctaatatgg ccgctcagat ccaatacttt gaagattcgt ccgtgggttat atggcagcat
 120
 gcggtggatg gtatcgtgta ccgaagtgcg gatgaaggca agtcgtgggc cccaattaag
 180
 gggcctgaac agggtcaggc gcaccttttc gtgctccatc cctacgacaa gactcaagcg
 240
 tatattctga cgcgcagcac tcagcattgg cgcacgtcga accgtggcga gacgtggcag
 300

tcattctcaa cgctcatcc gctacgacc ttgaaagcta tgctctgga ctttcatccg
 360
 acgcatcatg actggatcct tttcacgggc caggcttgca cggtaaa
 407

<210> 150
 <211> 135
 <212> PRT
 <213> Homo sapiens

<400> 150
 Xaa Leu Ala Trp Thr Leu Val Thr Gln Ala Ala Ile Pro Glu Val Lys
 1 5 10 15
 Val Thr His Phe Pro Asn Met Ala Ala Gln Ile Gln Tyr Phe Glu Asp
 20 25 30
 Ser Ser Val Val Ile Trp His Asp Ala Val Asp Gly Ile Val Tyr Arg
 35 40 45
 Ser Ala Asp Glu Gly Lys Ser Trp Ala Pro Ile Lys Gly Pro Glu Gln
 50 55 60
 Gly Gln Ala His Leu Phe Val Leu His Pro Tyr Asp Lys Thr Gln Ala
 65 70 75 80
 Tyr Ile Leu Thr Arg Ser Thr Gln His Trp Arg Thr Ser Asn Arg Gly
 85 90 95
 Glu Thr Trp Gln Ser Phe Ser Thr Pro His Pro Pro Thr Thr Leu Lys
 100 105 110
 Ala Met Pro Leu Asp Phe His Pro Thr His His Asp Trp Ile Leu Phe
 115 120 125
 Thr Gly Gln Ala Cys Thr Val
 130 135

<210> 151
 <211> 448
 <212> DNA
 <213> Homo sapiens

<400> 151
 accggtgtcc gtggctattg ccccgaaatgg tccccatccg cgtccccggg aactccctcg
 60
 gcttttcgcg catccaggtc cccagcccca gctactggtg cgccccgagc ccctaggtgc
 120
 cagagcgggtg gtcggccggg ctctgcecca gtctcggtc ctccctctc cccaccagaa
 180
 ggaaaaactt gggcccttcg agaaccctgt ggaatgttct ttgtaatcaa ctgtacatcc
 240
 gcttccacgg cacggcctcg tgcaaaatcg cgggtttcgg ggccctggag caaattgcgc
 300
 ttgtcagcgg cgacgtcagg aggacaaggg gaggggttcg cggctgaaac tgcagcttcg
 360
 cagcacagag ccatttttagg ctgctcccca cctcgcgggg cccatgggaa gccggccccg
 420
 ggagggcgcg gctgcatgga tattcgac
 448

<210> 152

<211> 149
 <212> PRT
 <213> Homo sapiens

<400> 152
 Thr Gly Val Arg Gly Tyr Cys Pro Glu Trp Ser Pro Ser Ala Ser Pro
 1 5 10 15
 Gly Thr Pro Ser Ala Phe Arg Ala Ser Arg Ser Pro Ala Pro Ala Thr
 20 25 30
 Gly Ala Pro Arg Ala Pro Arg Cys Gln Ser Gly Gly Arg Pro Gly Ser
 35 40 45
 Cys Pro Val Ser Ala Pro Pro Ser Ser Pro Pro Glu Gly Lys Thr Trp
 50 55 60
 Ala Leu Arg Glu Pro Cys Gly Met Phe Phe Val Ile Asn Cys Thr Ser
 65 70 75 80
 Ala Ser Thr Ala Arg Pro Arg Ala Lys Ser Arg Val Ser Gly Pro Trp
 85 90 95
 Ser Lys Leu Arg Leu Ser Ala Ala Thr Ser Gly Gly Gln Gly Glu Gly
 100 105 110
 Phe Ala Ala Glu Thr Ala Ala Ser Gln His Arg Ala Ile Leu Gly Cys
 115 120 125
 Ser Pro Pro Arg Gly Ala His Gly Lys Pro Ala Pro Gly Gly Arg Gly
 130 135 140
 Cys Met Asp Ile Arg
 145

<210> 153
 <211> 440
 <212> DNA
 <213> Homo sapiens

<400> 153
 nnntgggtcc atgtatgtgt gtgtatatga gggagacacg caggtgtgtg tccgagtgtg
 60
 tgtccatggg tccatgtatg tgtgtgtata tgtgggggaa caggtgtgtg tccgagtgtg
 120
 tgcattgggtc cgtgtatatg cgtgtatata tgcggggata tgtatatgtg tgtgtgtatg
 180
 aacaggtgta agtgggggagc actcaggtgt gtctgtgtgt gttcgtgtac acgtgtgtaa
 240
 gtgggtgacc atgaaggggt gtgtgtgtcc gtgtgtaggt ttgcgtgcat gcacacatgc
 300
 atgtgtgtac tggggcatcc aagcccctgg tctccactcc attccaccct acgcctacct
 360
 ccttgatctc tgcgcccagc cttggctgtg ctcccctgct gtagtcacgt ggggtgtctgc
 420
 acgtgggtgt ctgcacgcgt
 440

<210> 154
 <211> 69
 <212> PRT
 <213> Homo sapiens

<400> 154

Gly Arg His Ala Gly Val Cys Pro Ser Val Cys Pro Trp Val His Val
 1 5 10 15
 Cys Val Cys Ile Cys Gly Gly Thr Gly Val Cys Pro Ser Val Cys Met
 20 25 30
 Gly Pro Cys Ile Cys Val Tyr Ile Cys Gly Asp Met Tyr Met Cys Val
 35 40 45
 Cys Met Asn Arg Cys Lys Trp Gly Ala Leu Arg Cys Val Cys Val Cys
 50 55 60
 Ser Cys Thr Arg Val
 65

<210> 155

<211> 344

<212> DNA

<213> Homo sapiens

<400> 155

acgcgtatcg accaccatgt cgtcgtcacc acggcaagcg ctctcggcgg gcgagaacga
 60
 gtgaacatgg ccgagttgat ggccgatgcc gcgaccggca cgaaaccgtc ctacctacag
 120
 cgatcttctc cctcgatcac ctcgtttgaa gtggacaggg aacaaagaca ctcagacaac
 180
 gcgccgcagg aagtaaaaag ttcgctctcc gatcacggcc gtcgcgcgag tgcacagggg
 240
 gaactgggca cctcgcaagc tacgccaccg cgatccatgc ccccgcccgat atcttccgcc
 300
 tctctacct ccccttacc gatcagcatt atatccgac taga
 344

<210> 156

<211> 92

<212> PRT

<213> Homo sapiens

<400> 156

Met Ala Glu Leu Met Ala Asp Ala Ala Thr Gly Thr Lys Pro Ser Tyr
 1 5 10 15
 Leu Gln Arg Ser Ser Ser Ser Ile Thr Ser Phe Glu Val Asp Arg Glu
 20 25 30
 Gln Arg His Ser Asp Asn Ala Pro Gln Glu Val Lys Ser Ser Leu Ser
 35 40 45
 Asp His Gly Arg Arg Ala Ser Ala Gln Gly Glu Leu Gly Thr Ser Gln
 50 55 60
 Ala Thr Pro Pro Arg Ser Met Pro Pro Pro Val Ser Ser Ala Ser Ser
 65 70 75 80
 Thr Ser Pro Leu Pro Ile Ser Ile Ile Ser Asp Leu
 85 90

<210> 157

<211> 6816

<212> DNA

<213> Homo sapiens

<400> 157
nnagatctcc aaagaaccaa aaggatcaat atttctggat tcctgtatgg gtgtcgttca
60
gaacaacaaa gtcaggcggt ttgcttttga gctcaagatg caggacaaaa gtagttatct
120
cttggcagca gacagtgaag tggaaatgga agaattggatc acaattctaa ataagatect
180
ccagctcaac tttgaagctg caatgcaaga aaagcgaaat ggcgactctc acgaagatga
240
tgaacaaagc aaattggaag gttctgggtc cggttttagat agctacctgc cggaacttgc
300
caagagtgc agagaagcag aaatcaaact gaaaagtga agcagagtca aactttttta
360
tttggacca gatgccaga agcttgactt ctcatcagct gagccagaag tgaagtcatt
420
tgaagagaag tttggaaaaa ggatccttgt caagtgaat gatttatctt tcaatttgca
480
atgctgtgtt gccgaaaatg aagaaggacc cactacaaat gttgaacctt tctttgttac
540
tctatccctg tttgacataa aatacaaccg gaagatttct gccgatttcc acgtagacct
600
gaaccatttc tcagtggaggc aaatgatgcg caccacgtcc ccggcgctga tgaatggcag
660
tgggccgaaa cccaatctgc cctcaggggc atccttcag agccgccat gcagtatccg
720
aagcagggaa tattttcagt cacttgtcct catccagata tatttcttgt ggccagaatt
780
gaaaaagtcc ttcaggggag catcacacat tgcgctgagc catatatgaa aagttcagac
840
tcttctaagg tggccagaa ggtgctgaag aatgccaaagc aggcagcca aagactagga
900
cagtatagaa tgccatttgc ttgggcagca aggacattgt ttaaggatgc atctggaaat
960
cttgacaaaa atgccagatt ttctgccatc tacaggcaag acagcaataa gctatccaat
1020
gatgacatgc tcaagttact tgcagacttt cggaacacctg agaagatggc taagctccca
1080
gtgatttttag gcaatctaga cattacaatt gataatgttt cctcagactt ccctaattat
1140
gttaattcat catacattcc cacaaaaaa tttgaaacct gcagtaaaac tcccatcacg
1200
tttgaagtgg aggaatttgt gccctgcata ccaaaacaca ctcagcctta caccatctac
1260
accaatcacc tttacgttta tcctaagtac ttgaaatagc acagtcagaa gtcttttgc
1320
aaggctagaa atattgcgat ttgcattgaa ttcaaagatt cagatgagga agactctcag
1380
ccccttaagt gcatttatgg cagacctggg gggccagttt tcacaagaag cgcctttgct
1440
gcagttttac accatcacca aaaccagaa ttttatgatg agattaaaat agagttgccc
1500
actcagctgc atgaaaagca ccacctgttg ctcacattct tccatgtcag ctgtgacaac
1560

tcaagtaaag gaagcacgaa gaagagggat gtcgttgaaa cccaagttgg ctactcctgg
1620
cttccccctcc tgaaagacgg aagggtggtg acaagcgagc agcacatccc ggtctcggcg
1680
aaccttcctt cgggctatct tggctaccag gagcttggga tgggcaggca ttatggtccg
1740
gaaattaaat gggtagatgg aggcaagcca ctgctgaaaa tttccactca tctggtttct
1800
acagtgtata ctcaggatca gcatttacat aatTTTTTcc agtactgtca gaaaaccgaa
1860
tctggagccc aagccttagg aaacgaactt gtaaagtacc ttaagagtct gcatgcatg
1920
gaaggccacg tgatgatcgc cttcttgccc actatcctaa accagctgtt ccgagtcctc
1980
accagagcca cacaggaaga agtcgcggtt aacgtgactc gggtcattat tcatgtgggt
2040
gccagtgcc atgaggaagg attggagagc cacttgaggt catatgttaa gtacgcgtat
2100
aaggctgagc catatgttgc ctctgaatac aagacagtgc atgaagaact gaccaaatcc
2160
atgaccacga ttctcaagcc ttctgccgat ttcctcacca gcaacaaact actgaagtac
2220
tcatgggtttt tctttgatgt actgatcaaa tctatggctc agcatttgat agagaactcc
2280
aaagttaagt tgctgcgaaa ccagagattt cctgcacctc atcatcatgc agtggaaacc
2340
gttgtaaata tgctgatgcc acacatcact cagaagtttc gagataatcc agaggcatct
2400
aagaacgca atcatagcct tgctgtcttc atcaagagat gtttcacctt catggacagg
2460
ggctttgtct tcaagcagat caacaactac attagctgtt ttgctcctgg agacccaaag
2520
accctctttg aatacaagtt tgaatttctc cgtgtagtgt gcaaccatga acattatatt
2580
ccgttgaact taccaatgcc atttggaaaa ggcaggattc aaagatacca agacctccag
2640
cttgactact cattaacaga tgagttctgc agaaaccact tcttgggtggg actgttactg
2700
agggaggtgg ggacagccct ccaggagtgc cgggaggtcc gtctgatcgc catcagtgtg
2760
ctcaagaacc tgctgataaa gcattctttt gatgacagat atgcttcaag gagccatgag
2820
gcaaggatag ccacctcta cctgcctctg tttggtctgc tgattgaaaa cgtccagcgg
2880
atcaatgtga gggatgtgtc acccttcctt gtgaacgcgg gcatgactgt gaaggatgaa
2940
tccctggctc taccagctgt gaatccgctg gtgacgccgc agaaggggaag caccctggac
3000
aacagcctgc acaaggacct gctgggcgcc atctccggca ttgcttctcc atatacaacc
3060
tcaactccaa acatcaacag tgtgagaaat gctgattcga gaggatctct cataagcaca
3120
gattcgggta acagccttcc agaaaggaat agtgagaaga gcaattccct ggataagcac
3180

caacaaagta gcacattggg aaattccgtg gtccgctgtg ataaacttga ccagtctgag
3240
attaagagcc tactgatgtg tttcctctac atcttaaaga gcatgtctga tgatgctttg
3300
tttacatatt ggaacaaggc ttcaacatct gaacttatgg atttttttac aatatctgaa
3360
gtctgcctgc accagttcca gtacatgggg aagcgataca tagccagaac aggaatgatg
3420
catgccagat tgcagcagct gggcagcctg gataactctc tcacttttaa ccacagctat
3480
ggccactcgg acgcagatgt tctgcaccag tcattacttg aagccaacat tgctactgag
3540
gtttgcctga cagctctgga cacgctttct ctatttacat tggcgtttta gaaccagctc
3600
ctggccgacc atggacataa tcctctcatg aaaaaagttt ttgatgtcta cctgtgtttt
3660
cttcaaaaac atcagtctga aacggcttta aaaaatgtct tcactgcctt aaggtcctta
3720
atttataagt ttccctcaac attctatgaa gggagagcgg acatgtgtgc ggctctgtgt
3780
tacgagattc tcaagtgtg taactccaag ctgagctcca tcaggacgga ggctctccag
3840
ctgctctact tcctgatgag gaacaacttt gattacactg gaaagaagtc ctttgtccgg
3900
acacatttgc aagtcatcat atctgtcagc cagctgatag cagacgttgt tggcattggg
3960
ggaaccagat tccagcagtc cctgtccatc atcaacaact gtgccaacag tgaccggctt
4020
attaagcaca ccagcttctc ctctgatgtg aaggacttaa caaaaggat acgcacggtg
4080
ctaattggcca ccgcccagat gaaggagcat gagaacgacc cagagatgct ggtggacctc
4140
cagtacagcc tggccaaatc ctatgccagc acgcccagc tcaggaagac gtggctcgac
4200
agcatggcca ggatccatgt caaaaatggc gatctctcag aggcagcaat gtgctatgtc
4260
cacgtaacag ccctagtggc agaatatctc acacggaaaag aagcagtcca gtgggagccg
4320
ccccttctcc ccacagcca tagcgctgc ctgaggagga gccggggagg cgtgtttaga
4380
caaggatgca ccgccttcag ggtcattacc ccaaaccatg acgaggaggc ctccatgatg
4440
gaagacgtgg ggatgcagga tgtccatctc aacgaggatg tgctgatgga gtccttgag
4500
cagtgcgcag atggactctg gaaagccgag cgctacgagc tcattgccga catctacaaa
4560
cttatcatcc ccatttatga gaagcggagg gattttgaga ggctggccca tctgtatgac
4620
acgctgcacc gggcctacag caaagtgacc gaggtcatgc actcggggccg caggcttctg
4680
gggacctact tccgggtagc cttcttcggg caggcagcgc aataccagtt tacagacagt
4740
gaaacagatg tggagggatt ctttgaagat gaagatggaa aggagtatat ttacaaggaa
4800

cccaaactca caccgctgtc ggaaatttct cagagactcc ttaaactgta ctcgataaa
4860
tttggttctg aaaatgtcaa aatgatacag gattctggca aggtcaaccc taaggatctg
4920
gattctaagt atgcctacat ccaggtgact caggtcatcc ctttctttga cgaaaaagag
4980
ttgcaagaaa ggaaaacaga gtttgagaga tcccacaaca tccgcgctt catgtttgag
5040
atgccattta cgcagaccgg gaagaggcag ggcggggtgg aagagcagtg caaacggcgc
5100
accatcctga cagccataca ctgcttcctt tatgtgaaga agcgcacccc tgtcatgtac
5160
cagcaccaca ctgacctgaa ccccatcgag gtggccattg acgagatgag taagaagggtg
5220
gcggagctcc ggcagctgtg ctccctcgcc gaggtggaca tgatcaaact gcagctcaaa
5280
ctccagggca gcgtgagtgt tcaggtcaat gctggccac tagcatatgc gcgagctttc
5340
ttagatgata caaacacaaa gcgatatcct gacaataaag tgaagctgct taagggaagt
5400
ttcaggcaat ttgtggaagc ttgcggtcaa gccttagcgg taaacgaacg tctgattaaa
5460
gaagaccagc tcgagtatca ggaagaaatg aaagccaact acagggaaat ggcgaaggag
5520
ctttctgaaa tcatgcatga gcagatctgc cccctggagg agaagacgag cgtcttaccg
5580
aattcccttc acatcttcaa cgccatcagt gggactccaa caagcacaat ggttcacggg
5640
atgaccagct cgtcttcggt cgtgtgatta catctcatgg cccgtgtgtg gggacttgct
5700
ttgtcatttg caaactcagg atgctttcca aagccaatca ctggggagac cgagcacagg
5760
gaggaccaag gggaagggga gagaaaggaa ataaagaaca acgttatctt ttaacagact
5820
ttctatagga gttgtaagaa ggtgcacata tttttttaa tctcactggc aatattcaaa
5880
gttttcattg tgtcttaaca aagggtgtgt agacactctt gagctggact tagattttat
5940
tcttccttgc agagtagtgt tagaatagat ggccacaga aaaaaaagg tctgggatct
6000
acatggcagg gagggctgca ctgacattga tgccctgggg accttttgcc tcgaggctga
6060
gctggaaaat cttgaaaata tttttttttt cctgtggcac attcagggtg aatacaagaa
6120
ctatttttgt gactagtgtt tgatgacctt agggaaactga ccattgtaat ttttgtacca
6180
gtgaaccagg agatttagtg cttttatatt catttccttg catttaagaa aatatgaaag
6240
cttaaggaat tatgtgagct taaaactagt caagcagttt agaaccaaag gcctatatta
6300
ataaccgcaa ctatgctgaa aagtacaaag tagtacagta tattgttatg tacatatcat
6360
tgtaataca gtccctggcat tctgtacata tatgtattac atttctacat ttttaatact
6420

cacatgggct tatgcattaa gtttaattgt gataaatttg tgctgttcca gtatatgcaa
 6480
 tacacttttaa tggtttattc ttgtacataa aaatgtgcaa tatggagatg tatacagtct
 6540
 ttactatatt aggtttataa acagttttta gaatttcac cttttgccaa aatgggtggag
 6600
 tatgtaattg gtaaatacata aatcctgtgg tgaatgggtg tgtactttta agctgtcacc
 6660
 atgttatatt ttcttttaag acattaattt agtaatttta tatttgggaa aataaagggt
 6720
 tttaatttta tttaactgga atcaactgcc tgctgtaatt aaacattctg taccacatct
 6780
 gtattaaaaa gacattgctg accattaaaa aaaaaa
 6816

<210> 158
 <211> 1572
 <212> PRT
 <213> Homo sapiens

<400> 158
 Ala Ser Gly Asn Leu Asp Lys Asn Ala Arg Phe Ser Ala Ile Tyr Arg
 1 5 10 15
 Gln Asp Ser Asn Lys Leu Ser Asn Asp Met Leu Lys Leu Leu Ala
 20 25 30
 Asp Phe Arg Lys Pro Glu Lys Met Ala Lys Leu Pro Val Ile Leu Gly
 35 40 45
 Asn Leu Asp Ile Thr Ile Asp Asn Val Ser Ser Asp Phe Pro Asn Tyr
 50 55 60
 Val Asn Ser Ser Tyr Ile Pro Thr Lys Gln Phe Glu Thr Cys Ser Lys
 65 70 75 80
 Thr Pro Ile Thr Phe Glu Val Glu Glu Phe Val Pro Cys Ile Pro Lys
 85 90 95
 His Thr Gln Pro Tyr Thr Ile Tyr Thr Asn His Leu Tyr Val Tyr Pro
 100 105 110
 Lys Tyr Leu Lys Tyr Asp Ser Gln Lys Ser Phe Ala Lys Ala Arg Asn
 115 120 125
 Ile Ala Ile Cys Ile Glu Phe Lys Asp Ser Asp Glu Glu Asp Ser Gln
 130 135 140
 Pro Leu Lys Cys Ile Tyr Gly Arg Pro Gly Gly Pro Val Phe Thr Arg
 145 150 155 160
 Ser Ala Phe Ala Ala Val Leu His His His Gln Asn Pro Glu Phe Tyr
 165 170 175
 Asp Glu Ile Lys Ile Glu Leu Pro Thr Gln Leu His Glu Lys His His
 180 185 190
 Leu Leu Leu Thr Phe Phe His Val Ser Cys Asp Asn Ser Ser Lys Gly
 195 200 205
 Ser Thr Lys Lys Arg Asp Val Val Glu Thr Gln Val Gly Tyr Ser Trp
 210 215 220
 Leu Pro Leu Leu Lys Asp Gly Arg Val Val Thr Ser Glu Gln His Ile
 225 230 235 240
 Pro Val Ser Ala Asn Leu Pro Ser Gly Tyr Leu Gly Tyr Gln Glu Leu
 245 250 255
 Gly Met Gly Arg His Tyr Gly Pro Glu Ile Lys Trp Val Asp Gly Gly

			260					265					270			
Lys	Pro	Leu		Lys	Ile	Ser	Thr	His	Leu	Val	Ser	Thr	Val	Tyr	Thr	
		275					280					285				
Gln	Asp	Gln	His	Leu	His	Asn	Phe	Phe	Gln	Tyr	Cys	Gln	Lys	Thr	Glu	
	290					295					300					
Ser	Gly	Ala	Gln	Ala	Leu	Gly	Asn	Glu	Leu	Val	Lys	Tyr	Leu	Lys	Ser	
305					310					315					320	
Leu	His	Ala	Met	Glu	Gly	His	Val	Met	Ile	Ala	Phe	Leu	Pro	Thr	Ile	
			325					330						335		
Leu	Asn	Gln	Leu	Phe	Arg	Val	Leu	Thr	Arg	Ala	Thr	Gln	Glu	Glu	Val	
		340						345					350			
Ala	Val	Asn	Val	Thr	Arg	Val	Ile	His	Val	Val	Ala	Gln	Cys	His		
	355						360				365					
Glu	Glu	Gly	Leu	Glu	Ser	His	Leu	Arg	Ser	Tyr	Val	Lys	Tyr	Ala	Tyr	
	370					375					380					
Lys	Ala	Glu	Pro	Tyr	Val	Ala	Ser	Glu	Tyr	Lys	Thr	Val	His	Glu	Glu	
385					390					395					400	
Leu	Thr	Lys	Ser	Met	Thr	Thr	Ile	Leu	Lys	Pro	Ser	Ala	Asp	Phe	Leu	
			405					410						415		
Thr	Ser	Asn	Lys	Leu	Leu	Lys	Tyr	Ser	Trp	Phe	Phe	Phe	Asp	Val	Leu	
		420						425					430			
Ile	Lys	Ser	Met	Ala	Gln	His	Leu	Ile	Glu	Asn	Ser	Lys	Val	Lys	Leu	
	435						440					445				
Leu	Arg	Asn	Gln	Arg	Phe	Pro	Ala	Ser	Tyr	His	His	Ala	Val	Glu	Thr	
	450					455					460					
Val	Val	Asn	Met	Leu	Met	Pro	His	Ile	Thr	Gln	Lys	Phe	Arg	Asp	Asn	
465					470					475					480	
Pro	Glu	Ala	Ser	Lys	Asn	Ala	Asn	His	Ser	Leu	Ala	Val	Phe	Ile	Lys	
			485						490					495		
Arg	Cys	Phe	Thr	Phe	Met	Asp	Arg	Gly	Phe	Val	Phe	Lys	Gln	Ile	Asn	
		500						505					510			
Asn	Tyr	Ile	Ser	Cys	Phe	Ala	Pro	Gly	Asp	Pro	Lys	Thr	Leu	Phe	Glu	
	515						520					525				
Tyr	Lys	Phe	Glu	Phe	Leu	Arg	Val	Val	Cys	Asn	His	Glu	His	Tyr	Ile	
	530					535					540					
Pro	Leu	Asn	Leu	Pro	Met	Pro	Phe	Gly	Lys	Gly	Arg	Ile	Gln	Arg	Tyr	
545					550					555					560	
Gln	Asp	Leu	Gln	Leu	Asp	Tyr	Ser	Leu	Thr	Asp	Glu	Phe	Cys	Arg	Asn	
			565						570					575		
His	Phe	Leu	Val	Gly	Leu	Leu	Leu	Arg	Glu	Val	Gly	Thr	Ala	Leu	Gln	
		580						585					590			
Glu	Phe	Arg	Glu	Val	Arg	Leu	Ile	Ala	Ile	Ser	Val	Leu	Lys	Asn	Leu	

690 695 700
 Ser Thr Pro Asn Ile Asn Ser Val Arg Asn Ala Asp Ser Arg Gly Ser
 705 710 715 720
 Leu Ile Ser Thr Asp Ser Gly Asn Ser Leu Pro Glu Arg Asn Ser Glu
 725 730 735
 Lys Ser Asn Ser Leu Asp Lys His Gln Gln Ser Ser Thr Leu Gly Asn
 740 745 750
 Ser Val Val Arg Cys Asp Lys Leu Asp Gln Ser Glu Ile Lys Ser Leu
 755 760 765
 Leu Met Cys Phe Leu Tyr Ile Leu Lys Ser Met Ser Asp Asp Ala Leu
 770 775 780
 Phe Thr Tyr Trp Asn Lys Ala Ser Thr Ser Glu Leu Met Asp Phe Phe
 785 790 795 800
 Thr Ile Ser Glu Val Cys Leu His Gln Phe Gln Tyr Met Gly Lys Arg
 805 810 815
 Tyr Ile Ala Arg Thr Gly Met Met His Ala Arg Leu Gln Gln Leu Gly
 820 825 830
 Ser Leu Asp Asn Ser Leu Thr Phe Asn His Ser Tyr Gly His Ser Asp
 835 840 845
 Ala Asp Val Leu His Gln Ser Leu Leu Glu Ala Asn Ile Ala Thr Glu
 850 855 860
 Val Cys Leu Thr Ala Leu Asp Thr Leu Ser Leu Phe Thr Leu Ala Phe
 865 870 875 880
 Lys Asn Gln Leu Leu Ala Asp His Gly His Asn Pro Leu Met Lys Lys
 885 890 895
 Val Phe Asp Val Tyr Leu Cys Phe Leu Gln Lys His Gln Ser Glu Thr
 900 905 910
 Ala Leu Lys Asn Val Phe Thr Ala Leu Arg Ser Leu Ile Tyr Lys Phe
 915 920 925
 Pro Ser Thr Phe Tyr Glu Gly Arg Ala Asp Met Cys Ala Ala Leu Cys
 930 935 940
 Tyr Glu Ile Leu Lys Cys Cys Asn Ser Lys Leu Ser Ser Ile Arg Thr
 945 950 955 960
 Glu Ala Ser Gln Leu Leu Tyr Phe Leu Met Arg Asn Asn Phe Asp Tyr
 965 970 975
 Thr Gly Lys Lys Ser Phe Val Arg Thr His Leu Gln Val Ile Ile Ser
 980 985 990
 Val Ser Gln Leu Ile Ala Asp Val Val Gly Ile Gly Gly Thr Arg Phe
 995 1000 1005
 Gln Gln Ser Leu Ser Ile Ile Asn Asn Cys Ala Asn Ser Asp Arg Leu
 1010 1015 1020
 Ile Lys His Thr Ser Phe Ser Ser Asp Val Lys Asp Leu Thr Lys Arg
 1025 1030 1035 1040
 Ile Arg Thr Val Leu Met Ala Thr Ala Gln Met Lys Glu His Glu Asn
 1045 1050 1055
 Asp Pro Glu Met Leu Val Asp Leu Gln Tyr Ser Leu Ala Lys Ser Tyr
 1060 1065 1070
 Ala Ser Thr Pro Glu Leu Arg Lys Thr Trp Leu Asp Ser Met Ala Arg
 1075 1080 1085
 Ile His Val Lys Asn Gly Asp Leu Ser Glu Ala Ala Met Cys Tyr Val
 1090 1095 1100
 His Val Thr Ala Leu Val Ala Glu Tyr Leu Thr Arg Lys Glu Ala Val
 1105 1110 1115 1120
 Gln Trp Glu Pro Pro Leu Leu Pro His Ser His Ser Ala Cys Leu Arg

1125 1130 1135
 Arg Ser Arg Gly Gly Val Phe Arg Gln Gly Cys Thr Ala Phe Arg Val
 1140 1145 1150
 Ile Thr Pro Asn Ile Asp Glu Glu Ala Ser Met Met Glu Asp Val Gly
 1155 1160 1165
 Met Gln Asp Val His Phe Asn Glu Asp Val Leu Met Glu Leu Leu Glu
 1170 1175 1180
 Gln Cys Ala Asp Gly Leu Trp Lys Ala Glu Arg Tyr Glu Leu Ile Ala
 1185 1190 1195 1200
 Asp Ile Tyr Lys Leu Ile Ile Pro Ile Tyr Glu Lys Arg Arg Asp Phe
 1205 1210 1215
 Glu Arg Leu Ala His Leu Tyr Asp Thr Leu His Arg Ala Tyr Ser Lys
 1220 1225 1230
 Val Thr Glu Val Met His Ser Gly Arg Arg Leu Leu Gly Thr Tyr Phe
 1235 1240 1245
 Arg Val Ala Phe Phe Gly Gln Ala Ala Gln Tyr Gln Phe Thr Asp Ser
 1250 1255 1260
 Glu Thr Asp Val Glu Gly Phe Phe Glu Asp Glu Asp Gly Lys Glu Tyr
 1265 1270 1275 1280
 Ile Tyr Lys Glu Pro Lys Leu Thr Pro Leu Ser Glu Ile Ser Gln Arg
 1285 1290 1295
 Leu Leu Lys Leu Tyr Ser Asp Lys Phe Gly Ser Glu Asn Val Lys Met
 1300 1305 1310
 Ile Gln Asp Ser Gly Lys Val Asn Pro Lys Asp Leu Asp Ser Lys Tyr
 1315 1320 1325
 Ala Tyr Ile Gln Val Thr His Val Ile Pro Phe Phe Asp Glu Lys Glu
 1330 1335 1340
 Leu Gln Glu Arg Lys Thr Glu Phe Glu Arg Ser His Asn Ile Arg Arg
 1345 1350 1355 1360
 Phe Met Phe Glu Met Pro Phe Thr Gln Thr Gly Lys Arg Gln Gly Gly
 1365 1370 1375
 Val Glu Glu Gln Cys Lys Arg Arg Thr Ile Leu Thr Ala Ile His Cys
 1380 1385 1390
 Phe Pro Tyr Val Lys Lys Arg Ile Pro Val Met Tyr Gln His His Thr
 1395 1400 1405
 Asp Leu Asn Pro Ile Glu Val Ala Ile Asp Glu Met Ser Lys Lys Val
 1410 1415 1420
 Ala Glu Leu Arg Gln Leu Cys Ser Ser Ala Glu Val Asp Met Ile Lys
 1425 1430 1435 1440
 Leu Gln Leu Lys Leu Gln Gly Ser Val Ser Val Gln Val Asn Ala Gly
 1445 1450 1455
 Pro Leu Ala Tyr Ala Arg Ala Phe Leu Asp Asp Thr Asn Thr Lys Arg
 1460 1465 1470
 Tyr Pro Asp Asn Lys Val Lys Leu Leu Lys Glu Val Phe Arg Gln Phe
 1475 1480 1485
 Val Glu Ala Cys Gly Gln Ala Leu Ala Val Asn Glu Arg Leu Ile Lys
 1490 1495 1500
 Glu Asp Gln Leu Glu Tyr Gln Glu Glu Met Lys Ala Asn Tyr Arg Glu
 1505 1510 1515 1520
 Met Ala Lys Glu Leu Ser Glu Ile Met His Glu Gln Ile Cys Pro Leu
 1525 1530 1535
 Glu Glu Lys Thr Ser Val Leu Pro Asn Ser Leu His Ile Phe Asn Ala
 1540 1545 1550
 Ile Ser Gly Thr Pro Thr Ser Thr Met Val His Gly Met Thr Ser Ser

1555
Ser Ser Val Val
1570

1560

1565

<210> 159
<211> 540
<212> DNA
<213> Homo sapiens

<400> 159
gccggctctg ccatgtgctt actctgagcc acctaacctc ggctgcttc agtttactca
60
tccgctcatc tgcagaatgg gtgatgctgt cggctacttcg tggcatacag gaaagtgtccc
120
agcattgtca gcctcagtga gaggtggcca gtggggagtg gtggccactg tacacctggc
180
acagcccaga gatgcatgtg ccactctgtt gtgtgcttca accaaggggc gctctggcag
240
ggcttgggtg ggacttccca aagggcatgg aaaagttccc agtcaatgag atccatggag
300
acccatggga gtgggggtca gcccagcct aagaggaccc ccagccctgc cctgtgcccc
360
aggacacacc aggcactgtc ccttgtcgcc ttcccagaca acctgtaccc tccaggccac
420
cagttctcgt ccatgacaaa gaaaggagcc ttctaaataa gtgcccgcga gaggtgcac
480
gttccctgc cccttcggg tggacctggg tttcaaagag aagctgccag tgcaacgcgt
540

<210> 160
<211> 110
<212> PRT
<213> Homo sapiens

<400> 160
Met Val Ser Leu Ser Glu Arg Trp Pro Val Gly Ser Gly Gly His Cys
1 5 10 15
Thr Pro Gly Thr Ala Gln Arg Cys Met Cys His Ser Val Val Cys Phe
20 25 30
Asn Gln Gly Ala Leu Trp Gln Gly Leu Gly Gly Thr Ser Gln Arg Ala
35 40 45
Trp Lys Ser Ser Gln Ser Met Arg Ser Met Glu Thr His Gly Ser Gly
50 55 60
Gly Gln Pro Gln Pro Lys Arg Thr Pro Ser Pro Ala Leu Cys Pro Arg
65 70 75 80
Thr His Gln Ala Leu Ser Leu Val Ala Phe Pro Asp Asn Leu Tyr Pro
85 90 95
Pro Gly His Gln Phe Ser Ser Met Thr Lys Lys Gly Ala Phe
100 105 110

<210> 161
<211> 351
<212> DNA
<213> Homo sapiens

<400> 161
 nnacgcgtac gtctttcggc cgaagaagga acgtgggcag gggcctcctt cgctggccgc
 60
 cgcgcttggc tcgcagcgac gatgaagggc gacgacagca gcaagatcac ccacaagatc
 120
 gcccgggcca agcgcgaggg ccgcgtatgg tggagctttg agtacttccc gccgcgcacg
 180
 ccgcagggca tgcagaatct gtatgaccgt atcgagcgca tgagtcagct gggccccgag
 240
 tttgtggaca ttacgtggaa tgccgggggc cggacgtcgg atatgacgac gcagctggtc
 300
 aagacgggtgc atgcgtactt tgggtgctgag acgtgcatgc atctgacgtg c
 351

<210> 162
 <211> 117
 <212> PRT
 <213> Homo sapiens

<400> 162
 Xaa Arg Val Arg Leu Ser Ala Glu Glu Gly Thr Trp Ala Gly Ala Ser
 1 5 10 15
 Phe Ala Gly Arg Arg Ala Trp Leu Ala Ala Thr Met Lys Gly Asp Asp
 20 25 30
 Ser Ser Lys Ile Thr His Lys Ile Ala Arg Ala Lys Arg Glu Gly Arg
 35 40 45
 Val Trp Trp Ser Phe Glu Tyr Phe Pro Pro Arg Thr Pro Gln Gly Met
 50 55 60
 Gln Asn Leu Tyr Asp Arg Ile Glu Arg Met Ser Gln Leu Gly Pro Glu
 65 70 75 80
 Phe Val Asp Ile Thr Trp Asn Ala Gly Gly Arg Thr Ser Asp Met Thr
 85 90 95
 Thr Gln Leu Val Lys Thr Val His Ala Tyr Phe Gly Val Glu Thr Cys
 100 105 110
 Met His Leu Thr Cys
 115

<210> 163
 <211> 360
 <212> DNA
 <213> Homo sapiens

<400> 163
 gcgtgctcca tcggcacctt gcagatgggc gaattcgctg aaaacgtcgc cggtgccgctc
 60
 gacacctaca ccctgcgtca gcccatcggc gtatgcgcag gcatcactcc gttcaacttc
 120
 ccggcgatga ttccactgtg gatgttcccg atggcgattg cctgcggtaa cactttcgtg
 180
 ctcaaaccgt ccgaacaaga ccctctgtcg acgatgctgc tggtagaact ggcgctggaa
 240
 gccgggtgtgc cggccggcgt gctcaacgtg gtgcacggcg gcaaggatgt ggtggatgcg
 300

ctgtgcaccc ataaagatat caaggcagtt tctttcgtcg gttcgaccgc cgttggtacc
360

<210> 164

<211> 120

<212> PRT

<213> Homo sapiens

<400> 164

Ala	Cys	Ser	Ile	Gly	Thr	Leu	Gln	Met	Gly	Glu	Phe	Ala	Glu	Asn	Val
1				5					10				15		
Ala	Gly	Gly	Val	Asp	Thr	Tyr	Thr	Leu	Arg	Gln	Pro	Ile	Gly	Val	Cys
		20						25					30		
Ala	Gly	Ile	Thr	Pro	Phe	Asn	Phe	Pro	Ala	Met	Ile	Pro	Leu	Trp	Met
	35						40					45			
Phe	Pro	Met	Ala	Ile	Ala	Cys	Gly	Asn	Thr	Phe	Val	Leu	Lys	Pro	Ser
	50					55					60				
Glu	Gln	Asp	Pro	Leu	Ser	Thr	Met	Leu	Leu	Val	Glu	Leu	Ala	Leu	Glu
65					70					75				80	
Ala	Gly	Val	Pro	Ala	Gly	Val	Leu	Asn	Val	Val	His	Gly	Gly	Lys	Asp
				85					90					95	
Val	Val	Asp	Ala	Leu	Cys	Thr	His	Lys	Asp	Ile	Lys	Ala	Val	Ser	Phe
		100						105					110		
Val	Gly	Ser	Thr	Ala	Val	Gly	Thr								
		115					120								

<210> 165

<211> 728

<212> DNA

<213> Homo sapiens

<400> 165

gctagcagcc ttcaccctcc tagaggggca ggctcggcga caaggggcgg ggggtgccccg
60
tcccagcgag ggacgcccgg ggctgggggt gccggtcgag cccggggcaa cagcttcacc
120
aagtttgga accgcaacgt cttcatgaag gacaacagct cttcttcag cacagactcc
180
cgctcccgt cctcctccag gtccccgacg cgccacttcc gcagaagtga ctcccactca
240
gactccgaca gctcctactc aggggaatgag tgtcaccctg tgggccgcag gaaccgcgcc
300
cctaagggcc ggggcggtcg agggggcccat atggatcggg gccgaggcag ggcgcagcgt
360
gggaagaggc acgatctggc gcccaccaag cgcagtcgaa agaagatggc ggcgctggag
420
tgtgaggacc cggagcgaga gctgaagaag cagaagcggg cagcccgtt ccagcacgga
480
cactcccgc gctcgcct cgagcccctg gtgctgcaga tgagcagcct ggagagcagt
540
ggggctgacc ctgactggca ggagctgcag atcgtgggca cctgccctga catcaccaag
600
cactacctgc gctcacctg tgcccccgac ccgtccaccg tgcgccctgt ggcattccct
660

gtggcaggtt ttgaaaaagt cgctgtgcat ggtcaagtgc cactggaaag agaagcagga

720

ctacgcgt

728

<210> 166

<211> 242

<212> PRT

<213> Homo sapiens

<400> 166

Ala	Ser	Ser	Leu	His	Pro	Pro	Arg	Gly	Ala	Gly	Ser	Ala	Thr	Arg	Gly
1				5					10					15	
Gly	Gly	Ala	Pro	Ser	Gln	Arg	Gly	Thr	Pro	Gly	Ala	Gly	Gly	Ala	Gly
			20					25					30		
Arg	Ala	Arg	Gly	Asn	Ser	Phe	Thr	Lys	Phe	Gly	Asn	Arg	Asn	Val	Phe
			35				40				45				
Met	Lys	Asp	Asn	Ser	Ser	Ser	Ser	Ser	Thr	Asp	Ser	Arg	Ser	Arg	Ser
	50					55				60					
Ser	Ser	Arg	Ser	Pro	Thr	Arg	His	Phe	Arg	Arg	Ser	Asp	Ser	His	Ser
65					70				75					80	
Asp	Ser	Asp	Ser	Ser	Tyr	Ser	Gly	Asn	Glu	Cys	His	Pro	Val	Gly	Arg
				85				90						95	
Arg	Asn	Pro	Pro	Pro	Lys	Gly	Arg	Gly	Gly	Arg	Gly	Ala	His	Met	Asp
			100					105					110		
Arg	Gly	Arg	Gly	Arg	Ala	Gln	Arg	Gly	Lys	Arg	His	Asp	Leu	Ala	Pro
			115				120					125			
Thr	Lys	Arg	Ser	Arg	Lys	Lys	Met	Ala	Ala	Leu	Glu	Cys	Glu	Asp	Pro
	130					135					140				
Glu	Arg	Glu	Leu	Lys	Lys	Gln	Lys	Arg	Ala	Ala	Arg	Phe	Gln	His	Gly
145					150				155					160	
His	Ser	Arg	Arg	Leu	Arg	Leu	Glu	Pro	Leu	Val	Leu	Gln	Met	Ser	Ser
				165				170					175		
Leu	Glu	Ser	Ser	Gly	Ala	Asp	Pro	Asp	Trp	Gln	Glu	Leu	Gln	Ile	Val
			180					185					190		
Gly	Thr	Cys	Pro	Asp	Ile	Thr	Lys	His	Tyr	Leu	Arg	Leu	Thr	Cys	Ala
		195					200					205			
Pro	Asp	Pro	Ser	Thr	Val	Arg	Pro	Val	Ala	Phe	Pro	Val	Ala	Gly	Phe
	210					215					220				
Glu	Lys	Val	Ala	Val	His	Gly	Gln	Val	Pro	Leu	Glu	Arg	Glu	Ala	Gly
225					230					235				240	
Leu	Arg														

<210> 167

<211> 510

<212> DNA

<213> Homo sapiens

<400> 167

nnacgcgtgg aaccagaact caggcccgtg tgaggagtct ggtttgaac acacggggcc

60

gcaacacaga attgtcaggt cctgtgccgt gaccaccaac cctcggggcca tgccaggtgc

120

tgggtgagggg caggtggctc cgcagggcg cctgctggcc tgaccgcact cgtccacag
 180
 gtcctcatgg ggcctcctcg gctgggcttc gtgtccgcct acctctcaca gccactgctc
 240
 gatggctttg ccatgggggc ctccgtgacc atcctgacct cgcagctcaa acacctgctg
 300
 ggcgtgcgga tcccgcggca ccagggggcc ggcattggtg tcttcacatg gctgagcctg
 360
 ctgcgcggcg ccgggcaggc caacgtgtgc gacgtggtca ccagcacggg gtgcctggcg
 420
 gtgctgctag ccggaagga gctctcagac cgctaccgac accgcctgag ggtgccgctg
 480
 cccacggagc tgctggcat cgtgggtggc
 510

<210> 168
 <211> 128
 <212> PRT
 <213> Homo sapiens

<400> 168
 Gly Ala Gly Gly Ser Arg Gln Ala Pro Ala Gly Leu Thr Ala Leu Arg
 1 5 10 15
 Pro Gln Val Leu Met Gly Val Leu Arg Leu Gly Phe Val Ser Ala Tyr
 20 25 30
 Leu Ser Gln Pro Leu Leu Asp Gly Phe Ala Met Gly Ala Ser Val Thr
 35 40 45
 Ile Leu Thr Ser Gln Leu Lys His Leu Leu Gly Val Arg Ile Pro Arg
 50 55 60
 His Gln Gly Pro Gly Met Val Val Leu Thr Trp Leu Ser Leu Leu Arg
 65 70 75 80
 Gly Ala Gly Gln Ala Asn Val Cys Asp Val Val Thr Ser Thr Val Cys
 85 90 95
 Leu Ala Val Leu Leu Ala Ala Lys Glu Leu Ser Asp Arg Tyr Arg His
 100 105 110
 Arg Leu Arg Val Pro Leu Pro Thr Glu Leu Leu Val Ile Val Val Ala
 115 120 125

<210> 169
 <211> 537
 <212> DNA
 <213> Homo sapiens

<400> 169
 gaattccacc gcatgtcgtg tctggacgta tgtaggtcgc ggtagtgtgc gaccgccggt
 60
 gccttaaagg agagcgggca tcggcggtgc agtacgagag gggaagggtg gcggatactt
 120
 attgtcgggtg cggcatcgct catccacacc gtccgatggg tcaatggact ggtcaagcgg
 180
 ggtaacgagg ttcacctggc atcagtcctc ccggcgggcc gtcactccat tgatccccga
 240
 gttcggatcc acctggcccc acacggcggg aaggcaaaat acgtcgtcaa tgccggctgg
 300

ctgcgatcag tggcggctgg ggtgcaacct gacatcgtca acgtccacta tgcgaccggt
 360
 tatggctctgc tcgctcgtct tgcccatatt gacgccccga cgctgctgtc ggtgtgggga
 420
 agtgacgttt acgattcccc ccggggcaaatt cccctcatgc gtcacatggt ccgatccaac
 480
 ttggtctcag ctactcggat cgcacgcaca agccactgca tggcgcgtgt caccgct
 537

<210> 170

<211> 164

<212> PRT

<213> Homo sapiens

<400> 170

Cys	Ala	Thr	Ala	Gly	Ala	Leu	Lys	Glu	Ser	Gly	His	Arg	Arg	Cys	Ser
1				5					10					15	
Thr	Arg	Gly	Glu	Gly	Val	Arg	Ile	Leu	Ile	Val	Gly	Ala	Ala	Ser	Ser
		20						25					30		
Ile	His	Thr	Val	Arg	Trp	Val	Asn	Gly	Leu	Val	Lys	Arg	Gly	His	Glu
		35					40					45			
Val	His	Leu	Ala	Ser	Val	His	Pro	Ala	Gly	Arg	His	Ser	Ile	Asp	Pro
		50				55				60					
Arg	Val	Arg	Ile	His	Leu	Ala	Pro	His	Gly	Gly	Lys	Ala	Lys	Tyr	Val
65					70					75				80	
Val	Asn	Ala	Gly	Trp	Leu	Arg	Ser	Val	Ala	Ala	Gly	Val	Gln	Pro	Asp
			85					90					95		
Ile	Val	Asn	Val	His	Tyr	Ala	Thr	Gly	Tyr	Gly	Leu	Leu	Ala	Arg	Leu
		100						105					110		
Ala	His	Ile	Asp	Ala	Pro	Thr	Leu	Leu	Ser	Val	Trp	Gly	Ser	Asp	Val
		115					120					125			
Tyr	Asp	Ser	Pro	Arg	Ala	Asn	Pro	Leu	Met	Arg	His	Met	Val	Arg	Ser
		130				135					140				
Asn	Leu	Val	Ser	Ala	Thr	Arg	Ile	Ala	Ser	Thr	Ser	His	Cys	Met	Ala
145					150					155				160	
Arg	Val	Thr	Arg												

<210> 171

<211> 391

<212> DNA

<213> Homo sapiens

<400> 171

ctagacaagc tcgcgcgggt gggcttcgac actcttggtc tacagacctt cctaactgcg
 60
 ggggagaagg agtcccgcgc atggacgatt cacaagggcg acaccgcccc tgaggctgct
 120
 ggcgatcatcc ataccgactt ccagaagggg ttcacatcaagg cccaggtggt gtccttcggc
 180
 gaccttggtg aatttgccgg cgaaaaggag gccaggctg ctgggaagct gcggttgagg
 240
 ggcaaggagt acgttatgca ggacgggtgac gtagtggaat tccgatttaa cgtgtagctc
 300

cggtttgata cttacttggc ttaaccgcat ctgagatccg tcatatcttt ggcgtagcct
360

tattggtatg aataacatgc cgtagccaaa g
391

<210> 172

<211> 98

<212> PRT

<213> Homo sapiens

<400> 172

Leu	Asp	Lys	Leu	Ala	Arg	Val	Gly	Phe	Asp	Thr	Leu	Gly	Leu	Gln	Thr
1				5					10					15	
Phe	Leu	Thr	Ala	Gly	Glu	Lys	Glu	Ser	Arg	Ala	Trp	Thr	Ile	His	Lys
			20					25					30		
Gly	Asp	Thr	Ala	Pro	Glu	Ala	Ala	Gly	Val	Ile	His	Thr	Asp	Phe	Gln
		35				40					45				
Lys	Gly	Phe	Ile	Lys	Ala	Gln	Val	Val	Ser	Phe	Gly	Asp	Leu	Val	Glu
	50				55					60					
Phe	Gly	Gly	Glu	Lys	Glu	Ala	Gln	Ala	Ala	Gly	Lys	Leu	Arg	Leu	Glu
65				70				75						80	
Gly	Lys	Glu	Tyr	Val	Met	Gln	Asp	Gly	Asp	Val	Val	Glu	Phe	Arg	Phe
				85				90						95	

Asn Val

<210> 173

<211> 309

<212> DNA

<213> Homo sapiens

<400> 173

ccatggagtg tcccttgtgc gagcattttg agagctatac caacacccat cctgcaggt
60
cccagagccg agccatttct caggagagca ggaagggagc aggccgaggg gtgctcccag
120
ccagccccgg aacccgaggt ctggggacgc agccgaccag ccctccttgt ctgggcctct
180
gtttcctctt cgacacaggg aagcagggag gggccgatca gcgacttagg cctgttggct
240
gtggtgggggt cccctgcgtt tctgggaagc cacggaccct gggatgtacc tgggtttcat
300
tcgcagtga
309

<210> 174

<211> 102

<212> PRT

<213> Homo sapiens

<400> 174

Met	Glu	Cys	Pro	Leu	Cys	Glu	His	Phe	Glu	Ser	Tyr	Thr	Asn	Thr	His
.1				5				10					15		
Pro	Cys	Arg	Ser	Gln	Ser	Arg	Ala	Ile	Ser	Gln	Glu	Ser	Arg	Lys	Gly

[illegible]

```
<210> 175
<211> 8484
<212> DNA
<213> Homo sapiens
```

<400> 175					
nnacttttttt	tttttttttt	catttatgct	atggagaaac	cagcatggag	atgtcatggg
60					
agagcatgca	caggccccgc	cctagggagt	ggtgatgtgt	ttggggaggt	gcttgtttcc
120					
aggtccatcc	cacacgttgt	ccagttggat	cctatggcag	gctggctgtg	gctttctctc
180					
tcttgcttct	cttcctcctc	cagataaggg	tctgcaggat	cttctgctta	gcaagtgggtg
240					
gccaaggact	ggtggatggg	tggctggaag	cagcgacat	gctccacagt	ggaactgtct
300					
gtctccacgg	acttcatgta	tttgttcagg	atggcaaaaa	cctcattggt	caagatctga
360					
tacttctga	tccggtcggc	catcttcttc	aggggcacat	tcttaatgat	ttcatccttc
420					
ccgtcttgcc	tctgcacttt	tagcaggtgg	taacagaagt	cgaacaggtc	aaagcgacgc
480					
tgctggccca	gcaggacaat	gatggagcaa	ccagcccagt	tcaagccatc	gccgaaacac
540					
tgctcagctg	tgaactcggt	ggttcccaca	gggatgcagt	acacgaactg	catggcgctc
600					
cacagccggt	ggaactccac	acactcatcg	acgtgcatga	cgccattggt	gggcgggtggg
660					
ccccgccaga	tggggctctg	caggtagctc	cgaatgcggg	tcaggatgac	ctcaaacatg
720					
gacaggccac	agcacagccg	ctccttggtc	aggaggtcac	cctcgcgagc	aatggcgatt
780					
tgctgagggg	tccccagccg	ctcgatcaga	gggaccaggt	ggagcggggc	atacttggct
840					
tccagacgtt	tcattttggc	atcaagtctc	tccccctctt	tcacatggac	tcgcggcaag
900					
atgttctgga	aaggagccgc	gtgcagcagg	tcacacactt	cttctaaaga	cagagctctg
960					
ctcgatgagg	aggcagaaga	ggatggcatt	gcccacttcc	ctcaggctct	ggaagcacgt
1020					
ctgttttgag	ctctgcgtac	tcaatgatgt	ccttcagctg	gtggtggaag	aactccagga
1080					

tccctgggga gccatactca tgcgggggca agcggcatat cttgggcatc acctctatca
1140
gtgttttcac atactggaga atgggttcctt ggagcaagct cttcacaatc tttagcagtt
1200
cctccatgac cacagcgatg ccctgataac ccaggagtct gcagatagtc ttgaaagtgt
1260
ggaggtccca cgaagtcccg gtagctgccg taaatgctgg agtaggccaa gttcaaagcc
1320
ttggatccat gcagatactg aggctgtgca ttaggctgct tatctctttg aaattcctga
1380
gaaaatggta acactgtccg aacaaaccgg ttggtagagc cgttgtagca gtagttgggc
1440
aggaagtcac agttgagctc ccagaagacg tgcaggggtga tcctcccgtg gggcgctgac
1500
acgttggtgt tggcctcccg gaacatggcg tcgaagccgt ccagcgtcag gtaccggctc
1560
agcagcttgt gggctcatgcg gttgatttcc aacaggccat ccagctcaac tatggaggtc
1620
aaatcttcac tttcaaactg tccaatcgcc agttctaggg acttatacat ggctgctgag
1680
acgcgctggg tgatcagacg attgaggtct attgatctgc cgaggagctg cacatgcctc
1740
tgcttcagca gcgtctcgta gcggttagac ggcgggaggt ggatcgtggc tccctgattc
1800
ttgcattctg atcgtaaccg tttatcaaga agcaaacttc ctgccataac cttataatag
1860
gcaaatatct ggtctgccag cttgtagaca aactgatcaa aacacagggt cacctcagct
1920
tctatctcat cgtacaggaa ctgcttttta aacttgggtc gagcatagta ggcgctgtca
1980
ttgtacagat ccaggaggta gagcacgtac tccatcatgg aagggttcttt ggtttccagg
2040
atatgggtccg ttagaatcca gggcatcgac atctcaatgg ggaactggat cctcctgccc
2100
atggtcagct ccaggaagaa ctctcggaac cacagctgcg aaaggtcaca gactgctgc
2160
agcgtttcac tgaaatttat caagtgagtg tagaagaatg actctcgatg aaagtcctct
2220
atggcgagga caatgggtcc atccaggctg ctctcaggg tcttcttgga gccgcttttg
2280
tctgcaatga gtgattcaag catggttctc accatgtaca gctgtgtgct ggatggcccc
2340
acagcacgcc ggggcacctt gatatcaaat ccacctttgg ggtccttctc ccctctcaag
2400
catgggtcat tagggggctc tggccctccc tcccagtcac agatgggtct tcgaattgcc
2460
tgtaggacgc tgatgaggac attcttcttc ttccgtaccg cctgccgcag gggctcacgc
2520
agcgtcacct gggcgaagtc ctgcaatgcc gcgtagatgg tgttcctgat ggctgggtg
2580
aagacgctct ccatcctgcc catgagcacc tgcaggcctt tgatcatggc gatcacctca
2640
acgaaggcaa atttttctct actggtgtaa ttgtagcgtg tggtctctc atattcctcc
2700

gcggtgccag gacagtcctt gttgcagaac ttgtctgtgg gatgaaccag cttccaagag
2760
tacacctcca tgacgtgggc gctccacttg gatagaagct gcagaccccg cagggctagg
2820
tcgaagagct cgcgatactc ctctgtgac ttctggctgt ccagccctga gcccgtcacc
2880
acctcactgt tgctgtagcg agcgagctcg gagatgaagc ggatgtggtc atcccggatc
2940
tgaaccatct gctcgcagat attgtactgg gggctgatgc tgctctgggt gcacgtccat
3000
cgagatttat tttcctcgta gtgggcgtg gtcttgatat atcttgccag ttctatttgc
3060
atgtccccaa atagtggaaac cacctggagt tgcttgaagt acttgctgat tttggataag
3120
tttattcttt tcttggcatc caacttatag atgttactga cactcccatc catcaggtac
3180
agaccaaate ccatgacttt gagaagcatg tgtttctcac tgggcgtcaa atacatcctg
3240
ttctcgtagt aatccacaca cagattcaca atatctgcca ggagctcttc gtagccagaa
3300
atcacttcca gctgctgctg cagagactgt gtgatcttgt tatgattggc caggaacatg
3360
gacagattct gcgattcctg gatggactgt ggatctgcca ttttacgtaa aaactgtgct
3420
gccctcttgt aggcagagtg gtcgttcttg acgctgcact tcatgttctt cagatcgtcc
3480
agcacagcga acatgttgat gaatttggcc agtgtgatca ggtaggcttc tgacacgaag
3540
tccttcttgc gctcggcatg gcacagccgc ttcacctcgc tgcagaaccg ctcgatggcc
3600
ttgcgctgaa aatacatgaa cttcatgagc ttggtgacct ccggctccag cacctctact
3660
gtcttctcat agatctctac tcggttgggc tgctcgttgc atttcacctg gggaaatggcc
3720
cggaacagc tgcgccaggt gtacagcatg accgcatact catgtccttc ctccagcatc
3780
tcattcatgc tggagtggac tgtagcctgc tcaatgtacc ttgcaatgcc cgtgacaaat
3840
gcattcctgt cctcaaagtt tgtgtcaaag ttagcctggt acatgatgga ggaaggtgga
3900
ggctcgatgc atggctgctg gtcggggagg ggaagctctt caagcaggtc cacgttggac
3960
agggcatctt ccagggtgac gtgcgtggtc atggctgcag tttctgtatt ctgcactggc
4020
tctgcggcgg gtccagggga ggggcctgag cctaccgcag ggccccgcgc cggggcgag
4080
gatgctctgc cccgctccgc tgcgccgcca gtgcagaata cagaaactgc agccatgacc
4140
acgcacgtca cctggaaga tgccctgtcc aacgtggacc tgcttgaaga gcttccctc
4200
cccgaccagc agccatgcat cgagcctcca ccttctcca tcatgtacca ggctaacttt
4260
gacacaaact ttgaggacag gaatgcattt gtcacgggca ttgcaaggta cattgagcag
4320

gctacagtcc actccagcat gaatgagatg ctggagggaag gacatgagta tgcggtcatg
4380
ctgtacacct ggcgcagctg ctcccggggc atcccacagg tgaaatgtaa cgagcagcct
4440
aacagagtgg aaatctacga gaaaaccgtg gaggttcttg agcctgaggt cacaaaactg
4500
atgaatttca tgtacttcca gagaaatgcc attgagcgtt tctgcgggga agtgaggcgc
4560
ctgtgccatg ccgagaggag gaaggacttc gtgtcagaag cctacctgat cacactgggc
4620
aaattcatca acatgttcgc tgtgctggac gagctgaaga acatgaagtg cagtgtgaag
4680
aacgaccact cagcgtacaa gagggccgct cagtttttac gtaaaatggc agatccacag
4740
tccatccagg aatcgagaa tctgtccatg ttccctggcca atcataacaa gatcacacag
4800
tctctgcagc agcagctcga agtgatttct ggctacgaag agctcctggc agatattgtg
4860
aatctgtgtg tggattacta cgagaacagg atgtatttga cgcccagtga gaaacacatg
4920
cttctcaaag tcattgggatt tggctctgtac ctgatggatg ggagtgtcag taacatctat
4980
aagttggatg ccaagaaaag aataaactta tccaaaatcg acaagtactt caagcaactc
5040
caggtgggtc cgctatttgg ggacatgcaa atagaactgg caagatatat caagaccagc
5100
gccactacg aggaaaataa atctcgatgg acgtgcacat cctccggcag cagccctcag
5160
tacaacatct gcgagcagat gatccagatc cgcgaggacc acatgcgctt catttcggag
5220
ctggcgcgct acagcaacag cgaggtggtc acgggctcgg gccgccagga ggcccagaag
5280
acggacgcgg agtaccgcaa gctcttcgac ctggcgctgc agggcctgca gctgtgtcgc
5340
cagtgaggcg cgcacgtgat ggaagtgtat tcctggaagc ttgtgcaccc caccgacaag
5400
tactccaaca aggactgccc cgacagcgct gaagagtacg agcgtgccac gcgctacaac
5460
tacaccagcg aggagaagtt tgccctagtg gaggtgatcg ccatgatcaa aggcctgcag
5520
gtgctgatgg gcaggatgga gagcgtgttc aaccacgcca tccggcacac cgtctatgcc
5580
gcactgcagg acttctccca ggtgaccctt agggagccgc tgcggcaggc catcaagaag
5640
aagaagaacg tcatccagag tgtcctgcag gccatcagga agaccgtgtg tgactgggag
5700
acggggcatg agcccttcaa tgacccagcc ttgcggggcg agaaggaccc caagagcggc
5760
ttcgacataa aagtaccagc ccgcgccgtg ggaccctcca gcactcagct ttacatgggtg
5820
agaaccatgc tagagtccct cattgcagac aaaagtgggt ccaagaaaac cttgagaagt
5880
agccttgagg ggcccacat attggacata gaaaaatttc atcgagagtc attcttctac
5940

actcacttga taaatttcag tgaacgctg cagcagtgtg gtgaccttc gcagctgtgg
6000
ttccgagagt tcttcctgga gctgaccatg ggcaggagga tccagttccc cattgagatg
6060
tcgatgccct ggatcctgac ggaccacatc ctggagacca aggaggcatc gatgatggag
6120
tacgtgtctt actccctgga cctgtacaat gacagcgccc actacgcgct caccagggtc
6180
aacaagcagt tcctgtacga cgaaattgag gccgagggtga atctatgttt tgaccaat
6240
gtttacaagc tagcagacca gatatttgcc tattataagg ttatggcagg aagtttgctt
6300
cttgataaac ggttacgatc agaatgcaag aatcagggag ccacgatcca cctcccgccg
6360
tctaaccgct acgagacgct gctgaagcag aggcatgtgc agctcctcgg cagatcaata
6420
gacctcaatc gtctgatcac ccagcgctc tcagcagcca tgtataagtc cctagaactg
6480
gcgattggac gatttgaaag tgaagatttg acctccatag ttgagctgga tggcctgttg
6540
gaaatcaacc gcatgaccca caagctgctg agccgggtacc tgacgctgga cggcttcgac
6600
gccatgttcc gggaggccaa ccacaacgtg tcagcgccct acgggaggat caccctgcac
6660
gtcttctggg agctcaacta tgacttcctg cccaactact gctacaacgg ctctaccaac
6720
cggtttgttc ggacagtgtt accattttct caggaatttc aaagagataa gcagccta
6780
gcacagcctc agtatctgca tggatccaag gctttgaact tggcctactc cagcatttac
6840
ggcagctacc ggaacttcgt gggacctcca cactttcaag tcctctgccg gcttctcggc
6900
taccagggta tcgccgtggt catggaggag ctgctgaagg tcgtcaagag cctgctgcaa
6960
ggcacaatcc tgcagtacgt gaagacgctg atggagggtga tgcccaagat ctgccgcctg
7020
ccccggcacg agtacggctc tcctggtatc ctggagttct tccaccacca gctgaaggac
7080
atcgtggagt acgcagagct gaagacgggtg tgcttccaga acctgcggga ggtggggaac
7140
gccatcctct tetgctgct catcgagcag agcctgtctt tagaagaagt gtgtgacctg
7200
ctgcacgcgg ctctttcca gaacatcttg ccgcgagtc atgtgaaaga gggggagaga
7260
cttgatgcca aaatgaaaag actagaatca aagtacgccc ngctgcatct tgtccactg
7320
attgaaagac tggggacccc tcagcaaatt gccatcgcaa gagaggggga cctgctgaca
7380
aaggagcgcc tctgctgcgg cctgtccatg tttgaggtca tcctgacacg gatccggagc
7440
tttctggatg accccatctg gcgcgggcct ctgccagca atggggtcat gcatgtggac
7500
gagtgtgtgg agtttcacag actgtggagt gccatgcagt ttgtctactg cattcccg
7560

gggacacacg agttcacagt cgagcagtg c ttgggtgatg ggctacactg ggctggctgt
 7620
 atgatcatcg tacttcttgg gcagcagcgg cgttttgctg tgctggattt ctgctaccat
 7680
 ctacttaaag tccagaaaca tgatggcaaa gatgagatta ttaaaaatgt gcctttgaag
 7740
 aagatggtgg agagaattcg caagttccag attctcaatg atgagatcat caccatcctg
 7800
 gataagtacc tgaagtcagg cgacggggag ggcacgccag tggagcatgt gcgctgcttc
 7860
 cagccgcccc tccaccagtc cctcgccagc agctgagggc acgcgctgca ctccgtaact
 7920
 caacatggca tgcctttctc tccgtaaact atttagtgag atttttaggg actatttttc
 7980
 agtatctctg tacctgttaa agggggtgct ttctgatcta aaaacttaat ttataaaat
 8040
 tgacttattt ttctagacta aaattgtata tgcttttggg aattaggaac tcttgagaat
 8100
 attggctgct gattgttgcc atcacgttcc tacaaaattg tttttctatg ggatgttctg
 8160
 gcagctgtgt cataaaatgc tgctgggttc attcattcat tccataagaa acttaatacc
 8220
 agcaaatgca ttaaatccct tgccagttac cattaactgt aactatttag cttttgttta
 8280
 gggatctttc tgatggtctt ttatgagcaa tcttagttct aagtcattgt tcccatccct
 8340
 tttttgtgtg ttccagaaaa tagtgaactt gattcccctg cttccactaa atccagttgt
 8400
 gacaaaatct aacgtgacat cagatcgaaa ggttatagaa ataaaactaa tgagatctaa
 8460
 aaaaaaaaaa aaaaaaaaaa aaaa
 8484

<210> 176

<211> 1393

<212> PRT

<213> Homo sapiens

<400> 176

Met	Ala	Arg	Glu	Gln	Leu	Arg	Gln	Val	Tyr	Ser	Met	Thr	Ala	Tyr	Ser
1				5					10					15	
Cys	Pro	Ser	Ser	Ser	Ile	Ser	Phe	Met	Leu	Glu	Trp	Thr	Val	Ala	Cys
			20					25					30		
Ser	Met	Tyr	Leu	Ala	Met	Pro	Val	Thr	Asn	Ala	Phe	Leu	Ser	Ser	Lys
		35					40					45			
Phe	Val	Ser	Lys	Leu	Ala	Trp	Tyr	Met	Met	Glu	Glu	Gly	Gly	Gly	Ser
	50					55				60					
Met	His	Gly	Cys	Trp	Ser	Gly	Arg	Gly	Ser	Ser	Ser	Arg	Ser	Thr	
65				70					75				80		
Leu	Asp	Arg	Ala	Ser	Ser	Arg	Val	Thr	Cys	Val	Val	Met	Ala	Ala	Val
				85					90				95		
Ser	Val	Phe	Cys	Thr	Gly	Ser	Ala	Ala	Gly	Pro	Gly	Glu	Gly	Pro	Glu
			100					105					110		
Ala	Thr	Ala	Gly	Pro	Arg	Ala	Gly	Ala	Gln	Asp	Ala	Leu	Pro	Arg	Ser

115	120	125
Ala Ala Pro Pro Val Gln Asn Thr Glu Thr Ala Ala Met Thr Thr His		
130	135	140
Val Thr Leu Glu Asp Ala Leu Ser Asn Val Asp Leu Leu Glu Glu Leu		
145	150	155
Pro Leu Pro Asp Gln Gln Pro Cys Ile Glu Pro Pro Pro Ser Ser Ile		
165	170	175
Met Tyr Gln Ala Asn Phe Asp Thr Asn Phe Glu Asp Arg Asn Ala Phe		
180	185	190
Val Thr Gly Ile Ala Arg Tyr Ile Glu Gln Ala Thr Val His Ser Ser		
195	200	205
Met Asn Glu Met Leu Glu Glu Gly His Glu Tyr Ala Val Met Leu Tyr		
210	215	220
Thr Trp Arg Ser Cys Ser Arg Ala Ile Pro Gln Val Lys Cys Asn Glu		
225	230	235
Gln Pro Asn Arg Val Glu Ile Tyr Glu Lys Thr Val Glu Val Leu Glu		
245	250	255
Pro Glu Val Thr Lys Leu Met Asn Phe Met Tyr Phe Gln Arg Asn Ala		
260	265	270
Ile Glu Arg Phe Cys Gly Glu Val Arg Arg Leu Cys His Ala Glu Arg		
275	280	285
Arg Lys Asp Phe Val Ser Glu Ala Tyr Leu Ile Thr Leu Gly Lys Phe		
290	295	300
Ile Asn Met Phe Ala Val Leu Asp Glu Leu Lys Asn Met Lys Cys Ser		
305	310	315
Val Lys Asn Asp His Ser Ala Tyr Lys Arg Ala Ala Gln Phe Leu Arg		
325	330	335
Lys Met Ala Asp Pro Gln Ser Ile Gln Glu Ser Gln Asn Leu Ser Met		
340	345	350
Phe Leu Ala Asn His Asn Lys Ile Thr Gln Ser Leu Gln Gln Gln Leu		
355	360	365
Glu Val Ile Ser Gly Tyr Glu Glu Leu Leu Ala Asp Ile Val Asn Leu		
370	375	380
Cys Val Asp Tyr Tyr Glu Asn Arg Met Tyr Leu Thr Pro Ser Glu Lys		
385	390	395
His Met Leu Leu Lys Val Met Gly Phe Gly Leu Tyr Leu Met Asp Gly		
405	410	415
Ser Val Ser Asn Ile Tyr Lys Leu Asp Ala Lys Lys Arg Ile Asn Leu		
420	425	430
Ser Lys Ile Asp Lys Tyr Phe Lys Gln Leu Gln Val Val Pro Leu Phe		
435	440	445
Gly Asp Met Gln Ile Glu Leu Ala Arg Tyr Ile Lys Thr Ser Ala His		
450	455	460
Tyr Glu Glu Asn Lys Ser Arg Trp Thr Cys Thr Ser Ser Gly Ser Ser		
465	470	475
Pro Gln Tyr Asn Ile Cys Glu Gln Met Ile Gln Ile Arg Glu Asp His		
485	490	495
Met Arg Phe Ile Ser Glu Leu Ala Arg Tyr Ser Asn Ser Glu Val Val		
500	505	510
Thr Gly Ser Gly Arg Gln Glu Ala Gln Lys Thr Asp Ala Glu Tyr Arg		
515	520	525
Lys Leu Phe Asp Leu Ala Leu Gln Gly Leu Gln Leu Leu Ser Gln Trp		
530	535	540
Ser Ala His Val Met Glu Val Tyr Ser Trp Lys Leu Val His Pro Thr		

545 550 555 560
 Asp Lys Tyr Ser Asn Lys Asp Cys Pro Asp Ser Ala Glu Glu Tyr Glu
 565 570 575
 Arg Ala Thr Arg Tyr Asn Tyr Thr Ser Glu Glu Lys Phe Ala Leu Val
 580 585 590
 Glu Val Ile Ala Met Ile Lys Gly Leu Gln Val Leu Met Gly Arg Met
 595 600 605
 Glu Ser Val Phe Asn His Ala Ile Arg His Thr Val Tyr Ala Ala Leu
 610 615 620
 Gln Asp Phe Ser Gln Val Thr Leu Arg Glu Pro Leu Arg Gln Ala Ile
 625 630 635 640
 Lys Lys Lys Lys Asn Val Ile Gln Ser Val Leu Gln Ala Ile Arg Lys
 645 650 655
 Thr Val Cys Asp Trp Glu Thr Gly His Glu Pro Phe Asn Asp Pro Ala
 660 665 670
 Leu Arg Gly Glu Lys Asp Pro Lys Ser Gly Phe Asp Ile Lys Val Pro
 675 680 685
 Arg Arg Ala Val Gly Pro Ser Ser Thr Gln Leu Tyr Met Val Arg Thr
 690 695 700
 Met Leu Glu Ser Leu Ile Ala Asp Lys Ser Gly Ser Lys Lys Thr Leu
 705 710 715 720
 Arg Ser Ser Leu Glu Gly Pro Thr Ile Leu Asp Ile Glu Lys Phe His
 725 730 735
 Arg Glu Ser Phe Phe Tyr Thr His Leu Ile Asn Phe Ser Glu Thr Leu
 740 745 750
 Gln Gln Cys Asp Leu Ser Gln Leu Trp Phe Arg Glu Phe Phe Leu
 755 760 765
 Glu Leu Thr Met Gly Arg Arg Ile Gln Phe Pro Ile Glu Met Ser Met
 770 775 780
 Pro Trp Ile Leu Thr Asp His Ile Leu Glu Thr Lys Glu Ala Ser Met
 785 790 795 800
 Met Glu Tyr Val Leu Tyr Ser Leu Asp Leu Tyr Asn Asp Ser Ala His
 805 810 815
 Tyr Ala Leu Thr Arg Phe Asn Lys Gln Phe Leu Tyr Asp Glu Ile Glu
 820 825 830
 Ala Glu Val Asn Leu Cys Phe Asp Gln Phe Val Tyr Lys Leu Ala Asp
 835 840 845
 Gln Ile Phe Ala Tyr Tyr Lys Val Met Ala Gly Ser Leu Leu Leu Asp
 850 855 860
 Lys Arg Leu Arg Ser Glu Cys Lys Asn Gln Gly Ala Thr Ile His Leu
 865 870 875 880
 Pro Pro Ser Asn Arg Tyr Glu Thr Leu Leu Lys Gln Arg His Val Gln
 885 890 895
 Leu Leu Gly Arg Ser Ile Asp Leu Asn Arg Leu Ile Thr Gln Arg Val
 900 905 910
 Ser Ala Ala Met Tyr Lys Ser Leu Glu Leu Ala Ile Gly Arg Phe Glu
 915 920 925
 Ser Glu Asp Leu Thr Ser Ile Val Glu Leu Asp Gly Leu Leu Glu Ile
 930 935 940
 Asn Arg Met Thr His Lys Leu Leu Ser Arg Tyr Leu Thr Leu Asp Gly
 945 950 955 960
 Phe Asp Ala Met Phe Arg Glu Ala Asn His Asn Val Ser Ala Pro Tyr
 965 970 975
 Gly Arg Ile Thr Leu His Val Phe Trp Glu Leu Asn Tyr Asp Phe Leu

980 985 990
 Pro Asn Tyr Cys Tyr Asn Gly Ser Thr Asn Arg Phe Val Arg Thr Val
 995 1000 1005
 Leu Pro Phe Ser Gln Glu Phe Gln Arg Asp Lys Gln Pro Asn Ala Gln
 1010 1015 1020
 Pro Gln Tyr Leu His Gly Ser Lys Ala Leu Asn Leu Ala Tyr Ser Ser
 1025 1030 1035 1040
 Ile Tyr Gly Ser Tyr Arg Asn Phe Val Gly Pro Pro His Phe Gln Val
 1045 1050 1055
 Ile Cys Arg Leu Leu Gly Tyr Gln Gly Ile Ala Val Val Met Glu Glu
 1060 1065 1070
 Leu Leu Lys Val Val Lys Ser Leu Leu Gln Gly Thr Ile Leu Gln Tyr
 1075 1080 1085
 Val Lys Thr Leu Met Glu Val Met Pro Lys Ile Cys Arg Leu Pro Arg
 1090 1095 1100
 His Glu Tyr Gly Ser Pro Gly Ile Leu Glu Phe Phe His His Gln Leu
 1105 1110 1115 1120
 Lys Asp Ile Val Glu Tyr Ala Glu Leu Lys Thr Val Cys Phe Gln Asn
 1125 1130 1135
 Leu Arg Glu Val Gly Asn Ala Ile Leu Phe Cys Leu Leu Ile Glu Gln
 1140 1145 1150
 Ser Leu Ser Leu Glu Glu Val Cys Asp Leu Leu His Ala Ala Pro Phe
 1155 1160 1165
 Gln Asn Ile Leu Pro Arg Val His Val Lys Glu Gly Glu Arg Leu Asp
 1170 1175 1180
 Ala Lys Met Lys Arg Leu Glu Ser Lys Tyr Ala Pro Leu His Leu Val
 1185 1190 1195 1200
 Pro Leu Ile Glu Arg Leu Gly Thr Pro Gln Gln Ile Ala Ile Ala Arg
 1205 1210 1215
 Glu Gly Asp Leu Leu Thr Lys Glu Arg Leu Cys Cys Gly Leu Ser Met
 1220 1225 1230
 Phe Glu Val Ile Leu Thr Arg Ile Arg Ser Phe Leu Asp Asp Pro Ile
 1235 1240 1245
 Trp Arg Gly Pro Leu Pro Ser Asn Gly Val Met His Val Asp Glu Cys
 1250 1255 1260
 Val Glu Phe His Arg Leu Trp Ser Ala Met Gln Phe Val Tyr Cys Ile
 1265 1270 1275 1280
 Pro Val Gly Thr His Glu Phe Thr Val Glu Gln Cys Phe Gly Asp Gly
 1285 1290 1295
 Leu His Trp Ala Gly Cys Met Ile Ile Val Leu Leu Gly Gln Gln Arg
 1300 1305 1310
 Arg Phe Ala Val Leu Asp Phe Cys Tyr His Leu Leu Lys Val Gln Lys
 1315 1320 1325
 His Asp Gly Lys Asp Glu Ile Ile Lys Asn Val Pro Leu Lys Lys Met
 1330 1335 1340
 Val Glu Arg Ile Arg Lys Phe Gln Ile Leu Asn Asp Glu Ile Ile Thr
 1345 1350 1355 1360
 Ile Leu Asp Lys Tyr Leu Lys Ser Gly Asp Gly Glu Gly Thr Pro Val
 1365 1370 1375
 Glu His Val Arg Cys Phe Gln Pro Pro Ile His Gln Ser Leu Ala Ser
 1380 1385 1390
 Ser

<210> 177
 <211> 417
 <212> DNA
 <213> Homo sapiens

<400> 177
 acgcgtgatg tcacactgcc tctgccgctg ggtcctaatt cgattgcacg caccatggct
 60
 gcagttcgtg ggcgcgcatag tttctggcat gcttcgcgca tcctggagac cgatccccgc
 120
 gctgccgtga aaccgcctaa aaatgtgaag cgattgccc aagccgtgtc cgtggagcaa
 180
 atgcaaaagc tccttgccat acccagtctt aagactccta ccggcctgcg taatcgagcg
 240
 atacttgagt tcttatatgc taccggcgcg cgcgtagagc agatgctggc aacagacctg
 300
 gacgatatac acctgggcga aaaaccccg c gatgaaaacg gggaatctat tgcacttccc
 360
 gggatatgtg gcctttttgg aaaggaggt aaagagcgtt tagtcccttt gggatcc
 417

<210> 178
 <211> 139
 <212> PRT
 <213> Homo sapiens

<400> 178
 Thr Arg Asp Val Thr Leu Pro Leu Pro Leu Gly Pro Asn Ser Ile Ala
 1 5 10 15
 Arg Thr Met Ala Ala Val Arg Gly Ala His Ser Phe Trp His Ala Ser
 20 25 30
 Arg Ile Leu Glu Thr Asp Pro Ala Ala Ala Val Lys Pro Pro Lys Asn
 35 40 45
 Val Lys Arg Leu Pro Lys Ala Val Ser Val Glu Gln Met Gln Lys Leu
 50 55 60
 Leu Ala Ile Pro Ser Leu Lys Thr Pro Thr Gly Leu Arg Asn Arg Ala
 65 70 75 80
 Ile Leu Glu Phe Leu Tyr Ala Thr Gly Ala Arg Val Ser Glu Met Leu
 85 90 95
 Ala Thr Asp Leu Asp Asp Ile His Leu Gly Glu Lys Pro Arg Asp Glu
 100 105 110
 Asn Gly Glu Ser Ile Ala Leu Pro Gly Tyr Val Arg Leu Phe Gly Lys
 115 120 125
 Gly Gly Lys Glu Arg Leu Val Pro Leu Gly Ser
 130 135

<210> 179
 <211> 362
 <212> DNA
 <213> Homo sapiens

<400> 179
 acgcgtcgaa ggtgccggtg ggggcgatca ataacatcgc gcaatccctg gaagagcctc
 60

aggatgattgc ccgtgggttg atggtggaag atcccgcatc cccaagaatc cgggaattcg
 120
 ccattgggcc gggcagcccg aatccaaaat gtcggggcac gccagtgagg agtatggtaa
 180
 ggggcccggca ccgatgttgg nggcagcata cggatggaag tgctgggcga gcgcctgggt
 240
 ttgccggcag agcaactggg gcagctcaag gcgggcgggg tgatcgagca gttggattga
 300
 gcaatggcgg ccgcgaagcc cgccatttac cttgatgact gtttagcgcg cggattcttt
 360
 aa
 362

<210> 180
 <211> 108
 <212> PRT
 <213> Homo sapiens

<400> 180
 Met Ala Gly Phe Ala Ala Ala Ile Ala Gln Ser Asn Cys Ser Ile Thr
 1 5 10 15
 Pro Pro Ala Leu Ser Cys Pro Ser Cys Ser Ala Gly Lys Pro Arg Arg
 20 25 30
 Ser Pro Ser Thr Ser Ile Arg Met Leu Pro Pro Thr Ser Val Pro Ala
 35 40 45
 Pro Tyr His Thr Pro Thr Gly Arg Ala Pro Thr Phe Trp Ile Arg Ala
 50 55 60
 Ala Arg Pro Asn Gly Glu Phe Pro Asp Ser Trp Gly Cys Gly Ile Phe
 65 70 75 80
 His His Gln Pro Thr Gly Asn His Leu Arg Leu Phe Gln Gly Leu Arg
 85 90 95
 Asp Val Ile Asp Arg Pro His Arg His Leu Arg Arg
 100 105

<210> 181
 <211> 297
 <212> DNA
 <213> Homo sapiens

<400> 181
 gcgttgatca tgtccgaccc aggttgatc atgctggtac gccgtcactt cccgtgcatg
 60
 ccgattcact tgtcgttaca ggccaatacg gtgaattggg ccagcgtcga gttctggcaa
 120
 cagcaaggta tctgccgggt aatcctgtcg gggaattgt cactggaaga aatcggcgaa
 180
 atccgccaac aggtgccggc catggagctg gaagtgttg tgcacggcgc cctgtacatg
 240
 gcctattccg ggcgtgttt gttgtccggc tatatgaaca agcgcgatgc caaccaa
 297

<210> 182
 <211> 99
 <212> PRT

<213> Homo sapiens

<400> 182

Ala Leu Ile Met Ser Asp Pro Gly Leu Ile Met Leu Val Arg Arg His
 1 5 10 15
 Phe Pro Cys Met Pro Ile His Leu Ser Val Gln Ala Asn Thr Val Asn
 20 25 30
 Trp Ala Ser Val Glu Phe Trp Gln Gln Gln Gly Ile Cys Arg Val Ile
 35 40 45
 Leu Ser Arg Glu Leu Ser Leu Glu Glu Ile Gly Glu Ile Arg Gln Gln
 50 55 60
 Val Pro Ala Met Glu Leu Glu Val Phe Val His Gly Ala Leu Tyr Met
 65 70 75 80
 Ala Tyr Ser Gly Arg Cys Leu Leu Ser Gly Tyr Met Asn Lys Arg Asp
 85 90 95
 Ala Asn Gln

<210> 183

<211> 351

<212> DNA

<213> Homo sapiens

<400> 183

cgggacgtca ccatgaagcc gaccggctcg ggggatgtgg cgaacaaggt catcacccat
 60
 attccgttta acatcgtctc ccaggcgact catccattcc ttcgtacett ggacgatgtc
 120
 aagcgcattc ctttggcgac cgacgggctc ggccaccagg tcctgctcaa gggctaccag
 180
 gccgagggcc acgactacgc acaccccgac tacggcgcca acgtctccca ccgtgccggc
 240
 gggatgaagg atctcgagaa gtcaccgag tcgggcaggc agtgggaacac cgatttcggc
 300
 attcacgtca acctggtgga gtcctatcct gaggcgaatc acttcggcga c
 351

<210> 184

<211> 117

<212> PRT

<213> Homo sapiens

<400> 184

Arg Asp Val Thr Met Lys Pro Thr Gly Ser Gly Asp Val Ala Asn Lys
 1 5 10 15
 Val Ile Thr His Ile Pro Phe Asn Ile Val Ser Gln Ala Thr His Pro
 20 25 30
 Phe Leu Arg Thr Leu Asp Asp Val Lys Arg Ile Ser Leu Ala Thr Asp
 35 40 45
 Gly Leu Gly His Gln Val Leu Lys Gly Tyr Gln Ala Glu Gly His
 50 55 60
 Asp Tyr Ala His Pro Asp Tyr Gly Gly Asn Val Ser His Arg Ala Gly
 65 70 75 80
 Gly Met Lys Asp Leu Glu Lys Leu Thr Glu Ser Gly Arg Gln Trp Asn

85 90 95
 Thr Asp Phe Gly Ile His Val Asn Leu Val Glu Ser Tyr Pro Glu Ala
 100 105 110
 Asn His Phe Gly Asp
 115

<210> 185
 <211> 396
 <212> DNA
 <213> Homo sapiens

<400> 185
 cgcggtgggtc tcagtaaaga aaatttggtg cttagaggat gcaccattag aaacacagag
 60
 gctgttggtg gcattgtggt ttatgcaggc catgaaacca aagcaatgct gaacaacagt
 120
 gggccacggt ataagcgcag caaattagaa agaagagcaa acacagatgt cctctggtgt
 180
 gtcattgttc tggtcataat gtgcttaact ggcgcagtag gtcattggaat ctggctgagc
 240
 aggtatgaaa agatgcattt tttcaatggt cccgagcctg atggacatat catatcacca
 300
 ctgttggcag gattttatat gttttggacc gtgatcattt tgttacaggt cttgattcct
 360
 atttctctct atgtttccat cgaaattgtg aagctt
 396

<210> 186
 <211> 132
 <212> PRT
 <213> Homo sapiens

<400> 186
 Arg Val Gly Leu Ser Lys Glu Asn Leu Leu Arg Gly Cys Thr Ile
 1 5 10 15
 Arg Asn Thr Glu Ala Val Val Gly Ile Val Val Tyr Ala Gly His Glu
 20 25 30
 Thr Lys Ala Met Leu Asn Asn Ser Gly Pro Arg Tyr Lys Arg Ser Lys
 35 40 45
 Leu Glu Arg Arg Ala Asn Thr Asp Val Leu Trp Cys Val Met Leu Leu
 50 55 60
 Val Ile Met Cys Leu Thr Gly Ala Val Gly His Gly Ile Trp Leu Ser
 65 70 75 80
 Arg Tyr Glu Lys Met His Phe Phe Asn Val Pro Glu Pro Asp Gly His
 85 90 95
 Ile Ile Ser Pro Leu Leu Ala Gly Phe Tyr Met Phe Trp Thr Val Ile
 100 105 110
 Ile Leu Leu Gln Val Leu Ile Pro Ile Ser Leu Tyr Val Ser Ile Glu
 115 120 125
 Ile Val Lys Leu
 130

<210> 187
 <211> 423

<212> DNA

<213> Homo sapiens

<400> 187

cgagtgtca ccgcgtcagc cgtcatgcgt cccactgagg ctgttgcttc tcggtcggca
60
gaacctcgac gagttcagcg gatcctggac cagcgcgagt gggctggcgt ctctgttctc
120
gatgagcatc gtcgtttgct tggcacggtc ggcgatcaag aggtcatcga ggctgctcgc
180
cgcgagatc gcagtattgc tgacgcgggtg gaaactaacg gcattctcac ggcgcggacc
240
gacactccgt tgtccgagct ctctgctccg accagcaacg ccaggggtgcc gttggccgct
300
gtcgacgagg acttccacct catgggtgtc atctctcggg tgacctgct cgacgcgatg
360
tcacgagctc gcgacgaggc aggagaggga tctgtcatgt ccttgagaa caccggaaag
420
ctt
423

<210> 188

<211> 141

<212> PRT

<213> Homo sapiens

<400> 188

Arg Val Leu Thr Ala Ser Ala Val Met Arg Pro Thr Glu Ala Val Val
1 5 10 15
Ser Arg Ser Ala Glu Pro Arg Arg Val Gln Arg Ile Leu Asp Gln Arg
20 25 30
Glu Trp Ala Gly Val Phe Val Val Asp Glu His Arg Arg Leu Leu Gly
35 40 45
Thr Val Gly Asp Gln Glu Val Ile Glu Ala Ala Arg Arg Gly Asp Arg
50 55 60
Ser Ile Ala Asp Ala Val Glu Thr Asn Gly Ile Leu Thr Ala Arg Thr
65 70 75 80
Asp Thr Pro Leu Ser Glu Leu Phe Ala Pro Thr Ser Asn Ala Arg Val
85 90 95
Pro Leu Ala Val Val Asp Glu Asp Phe His Leu Met Gly Val Ile Ser
100 105 110
Arg Val Thr Leu Leu Asp Ala Met Ser Arg Ala Arg Asp Glu Ala Gly
115 120 125
Glu Gly Ser Val Met Ser Leu Glu Asn Thr Gly Lys Leu
130 135 140

<210> 189

<211> 429

<212> DNA

<213> Homo sapiens

<400> 189

ngatgggtta ccaacatatg cagggttcga gcggcaatag ctctcgggg gctggcagtg
60

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.